

Avec les Nuls, tout devient facile !

L'Intelligence artificielle pour les nuls



- Le rôle de l'IA dans notre société
- Comment l'IA fait fonctionner les robots, les drones et les voitures sans conducteur
- Les limites de ce que l'IA peut faire

John Paul Mueller
Luca Massaron



L’Intelligence artificielle pour les nuls

John Paul Mueller
et Luca Massaron



L'Intelligence artificielle pour les Nuls

Titre de l'édition originale : ***Artificial Intelligence For Dummies***
Pour les Nuls est une marque déposée de Wiley Publishing, Inc.
For Dummies est une marque déposée de Wiley Publishing, Inc.

Edition française publiée en accord avec Wiley Publishing, Inc.

Éditions First, un département d'Édi8
12 avenue d'Italie
75013 Paris
Tél. : 01 44 16 09 00
Fax : 01 44 16 09 01
E-mail : firstinfo@efirst.com
Web : www.editionsfirst.fr

ISBN : 978-2-412-04352-3
ISBN numérique : 978-2-412-04903-7
Dépôt légal : 1er trimestre 2019

Collection dirigée par Jean-Pierre Cano
Traduction : Marc Rozenbaum
Mise en page : Marie Housseau

Ce livre numérique a été converti initialement au format EPUB par Isako www.isako.com à partir de l'édition papier du même ouvrage.

Introduction

À coup sûr, vous avez déjà eu de multiples occasions d'entendre parler de l'intelligence artificielle (en abrégé : IA). Elle est aujourd'hui présente dans les films, dans les livres, dans les journaux et, bien entendu, sur l'Internet. Elle intervient dans le fonctionnement des robots, des voitures sans conducteur, des drones, des équipements médicaux, des sites de vente en ligne, et dans toutes sortes de technologies susceptibles d'influencer notre quotidien de diverses manières.

Souvent, les commentateurs et les spécialistes vous inondent d'informations (plus ou moins justes) concernant l'IA. Pour certains, l'IA, ce sont de drôles de petites prouesses technologiques, tandis que selon d'autres, l'IA représenterait une sombre menace pour la survie de l'humanité. Le problème que pose une telle quantité d'informations disparates, c'est qu'il devient bien difficile de faire le tri entre ce qui est réel et ce qui n'est que le produit d'une imagination débridée. Toute cette frénésie autour de l'IA est en grande partie le résultat d'attentes démesurées et irréalistes de la part des scientifiques, des entrepreneurs et autres professionnels. Si vous avez l'impression de ne vraiment rien savoir sur cette technologie, qui semble pourtant être devenue un élément essentiel de votre vie quotidienne, *L'Intelligence artificielle pour les Nuls* est le livre qu'il vous faut.

En consultant les médias, vous pourrez remarquer que la plupart du temps, les technologies les plus utiles sont assez rébarbatives. Ce n'est certainement pas le genre de

discipline qui suscite un enthousiasme délirant. Il en est ainsi de l'IA : elle est assez répandue pour être devenue routinière. Vous-même, vous vous en servez sous une forme ou une autre, mais sans même vous en rendre compte tant il s'agit d'une chose banale. Ce livre vous sensibilisera à ces utilisations très réelles et essentielles de l'IA. Un petit thermostat pour réguler la température chez vous n'a peut-être pas grand-chose de passionnant en soi, mais c'est pourtant un exemple extraordinairement pratique d'application d'une technologie qui effraye un certain nombre de gens.

Bien sûr, ce livre traite aussi d'applications très sympathiques de l'IA. Peut-être ne savez-vous pas, par exemple, qu'il existe aujourd'hui un dispositif de surveillance médicale permettant de prédire le délai dans lequel un problème cardiaque peut survenir. Oui, ce dispositif existe. L'IA fait aussi fonctionner les drones et les automobiles, et elle rend possibles toutes sortes de robots. On l'utilise aujourd'hui dans toutes sortes d'applications spatiales, et elle jouera un rôle prédominant dans toutes les aventures spatiales de demain.

Contrairement à bien d'autres ouvrages sur ce sujet, *L'Intelligence artificielle pour les Nuls* vous révèle aussi la vérité sur les limites que rencontrent les applications de l'IA. Ainsi, il est certaines activités et tâches essentielles que l'IA ne pourra jamais prendre en charge, et il en est d'autres qu'elle ne pourra pas prendre en charge avant longtemps. Certains vous affirmeront le contraire, mais ce livre vous explique pourquoi cela ne peut pas fonctionner, et il fait table rase de tout ce battage qui n'a pu que vous laisser dans l'ignorance concernant l'IA. Une des idées essentielles, dans ce livre, est que l'être humain sera toujours important. On peut même affirmer qu'avec l'IA, l'être humain devient encore plus important, sachant que l'IA est ce qui permet aux humains d'exceller dans des

domaines auxquels vous n’auriez sans doute jamais pensé.

À propos de ce livre

L’Intelligence artificielle pour les Nuls vous permet tout d’abord de comprendre ce qu’est l’IA, et plus particulièrement, ce dont elle a besoin pour fonctionner et pourquoi elle a été un échec dans le passé. Ce livre vous fournit aussi les bases concernant les problèmes que pose actuellement l’IA et il vous explique pourquoi, dans certains cas, ces problèmes risquent d’être pratiquement impossibles à résoudre. Bien entendu, vous allez aussi découvrir les solutions à certains problèmes et les pistes que suivent les scientifiques pour trouver des réponses en matière d’IA.

Pour qu’une technologie survive, il faut qu’elle débouche sur un ensemble d’applications concrètes, viables et efficaces. Elle doit aussi être rentable pour les investisseurs. Si l’IA n’a pas eu le succès attendu dans le passé, c’est parce que toutes ces conditions n’étaient pas réunies. L’IA a aussi souffert d’être en avance sur son temps : pour pouvoir vraiment aboutir, l’IA avait besoin des matériels informatiques dont nous disposons actuellement. Aujourd’hui, l’IA est utilisée dans diverses applications informatiques et dans l’automatisation des processus. Elle joue aussi un rôle essentiel dans le domaine médical et dans l’amélioration des interactions humaines. L’IA intervient également dans l’analyse des données, dans l’apprentissage machine et dans l’apprentissage profond. Ces termes peuvent parfois prêter à confusion. C’est pourquoi une des raisons pour lesquelles il est conseillé de lire ce livre est qu’ainsi, vous découvrirez de quelle façon ces technologies sont liées entre elles.

Aujourd'hui, l'avenir de l'IA est véritablement prometteur, car elle est devenue une technologie essentielle. Ce livre vous indique aussi les orientations que l'IA devrait suivre dans l'avenir. Les différentes tendances envisagées dans ce livre sont déduites de ce que l'on essaye effectivement de faire aujourd'hui avec l'intelligence artificielle. C'est une nouvelle technologie qui n'a pas encore décollé, mais on y travaille, si bien qu'elle a de bonnes chances de faire ses preuves à un moment ou à un autre.

Pour que vous puissiez encore plus facilement assimiler les concepts, ce livre utilise les conventions suivantes :

» Les adresses Internet apparaissent en `police courrier`. Si vous lisez une version numérique de ce livre sur un appareil relié à l'Internet, vous pouvez cliquer sur l'adresse Internet suivante pour visiter le site Web qui lui correspond : `www.pourlesnuls.fr`.

» Les mots en *italique* sont des termes particuliers qu'il convient de connaître. Ils sont utilisés (parfois à mauvais escient) de différentes manières dans la presse et dans d'autres médias comme les films. Connaître leur signification vous permettra d'y voir plus clair en ce qui concerne l'IA.

Icônes utilisées dans ce livre

Au cours de votre lecture, vous verrez dans la marge ces icônes qui signalent des points intéressants (ou non, selon le cas). Cette section présente brièvement chaque icône.



Cette icône vous indique des techniques qui vous permettront d'économiser du temps ou d'accomplir une tâche en évitant un surcroît de travail, ou bien, des ressources à consulter pour tirer le meilleur profit de votre initiation à l'IA.



Nous ne souhaitons pas apparaître comme des parents mécontents ni comme des maniaques, mais nous vous conseillons vivement d'éviter tout ce qui est signalé par cette icône. Sans quoi, vous risqueriez d'être induit en erreur, comme tous ces gens qui semblent avoir très peur de l'IA.



Cette icône indique une innovation ou une avancée technique. Ce genre d'information succincte peut parfois vous être utile. Il se peut qu'elle soit la solution dont vous avez besoin dans le cadre d'une application donnée. Ce sont des passages que vous pouvez aussi sauter.



Si vous ne devez rien retenir d'autre du contenu d'un chapitre ou d'une section, retenez au moins l'information mise en évidence par cette icône. Il s'agira généralement d'un processus essentiel, ou bien d'une chose à savoir pour pouvoir interagir avec l'IA de façon satisfaisante.

Pour aller plus loin

Votre découverte de l'IA ne se limitera pas à la lecture de ce livre, qui n'en est en réalité que le commencement. Nous vous indiquons des ressources en ligne pour rendre cette lecture plus flexible, et pour que ce livre réponde mieux à vos besoins. Si vous envoyez un e-mail à John Mueller, nous serons en mesure de répondre à vos questions et de vous préciser quel impact les dernières nouveautés dans le domaine de l'IA et des technologies connexes peuvent avoir sur le contenu de ce livre. Vous pourrez accéder à tous ces suppléments intéressants :

- **Aide-mémoire :** les antisèches, au lycée, pour avoir une meilleure note, vous vous souvenez ? Ici, c'est un peu la même chose. Ce « *cheat sheet* » est constitué de notes concernant des travaux avec l'IA que tout le monde ne connaît pas. Vous trouverez un aide-mémoire associé à ce livre en vous rendant sur

le site www.dummies.com. L'aide-mémoire est constitué d'informations claires, comme la signification de tous ces étranges acronymes et abréviations associés à l'IA, à l'apprentissage machine et à l'apprentissage profond.

- **Mises à jour :** parfois, les choses évoluent. Il se peut que nous n'ayons pas anticipé un changement imminent quand nous avons rédigé ce livre, bien que nous ayons consulté des boules de cristal. Auparavant, on considérait simplement que le livre cessait d'être très utile car son contenu était périmé, mais vous avez désormais la possibilité de trouver en ligne des mises à jour de votre livre, en vous rendant sur le site www.dummies.com et en effectuant une recherche sur le titre de ce livre (en anglais : *Artificial Intelligence For Dummies*).
- Outre ces mises à jour, consultez le blog de l'auteur. Vous y trouverez des réponses aux questions des lecteurs et des démonstrations de techniques utiles. L'adresse du blog est http://blog.johnmuellerbooks.com/.

Par où commencer ?

Il est temps pour vous de commencer à découvrir l'IA et ce qu'elle peut vous apporter. Vous n'êtes pas obligé de lire tous les chapitres du livre, mais si vous êtes novice, il est conseillé de commencer par le Chapitre 1. Cela vous permettra d'assimiler les bases dont vous aurez besoin pour que la lecture du reste du livre vous soit profitable.

Si votre principal objectif, en lisant ce livre, est de savoir dans quels domaines on utilise l'IA aujourd'hui, l'information se trouve dans la Deuxième partie. Rendez-vous au Chapitre 5.

Le lecteur qui possède déjà une connaissance plus avancée de l'IA pourra prendre pour point de départ le Chapitre 9. La Troisième partie de ce livre présente tout ce qu'il y a de plus récent dans ce domaine. Si vous ne vous souciez pas de savoir comment l'IA est utilisée au niveau élémentaire (non pas par les développeurs, mais simplement par les profanes qui s'y intéressent), vous préférerez peut-être sauter une partie du livre.

Vous voulez connaître les utilisations les plus incroyables de l'IA, actuelles et futures ? Dans ce cas, consultez directement le Chapitre 12. La Quatrième partie et la Cinquième partie vous expliquent tout cela, sans que vous soyez obligé pour autant d'ingurgiter une énorme masse d'informations. Les informations contenues dans la Quatrième partie sont consacrées aux matériels dont le fonctionnement repose sur l'IA, tandis que la Cinquième partie traite plutôt des débouchés futurs de l'IA.

PARTIE 1
Introduction à l'IA

DANS CETTE PARTIE…

» Découvrir ce que l'IA peut réellement apporter

» Prendre la mesure de l'incidence des données sur l'utilisation de l'IA

» Comprendre comment l'IA utilise des algorithmes pour effectuer des tâches

» Considérer l'utilisation d'un matériel spécialisé pour une meilleure efficacité de l'IA

Chapitre 1

Introduction à l’IA

DANS CE CHAPITRE…

» **Définir l’IA et résumer son historique**

» **Utiliser l’IA pour des tâches concrètes**

» **Faire le tri dans la masse d’informations sur l’IA**

» **Cerner le lien entre l’IA et l’informatique**

Une des raisons pour lesquelles l’intelligence artificielle (IA) a connu plusieurs faux départs au cours du temps est que les gens ne comprennent pas vraiment en quoi elle consiste, ni même ce que doit être son rôle. Le problème est que le cinéma, la télévision et la littérature ont donné au public de faux espoirs dans ce domaine. La tendance répandue à l’anthropomorphisme (le fait d’attribuer des caractéristiques humaines à des êtres ou à des choses) a également contribué à faire accroire que l’IA allait accomplir davantage que ce qu’il était raisonnable d’en attendre. Pour commencer, le mieux est donc de définir ce qu’est vraiment l’IA, ce qu’elle n’est pas, et son lien actuel avec l’informatique.



Naturellement, ce que l’on peut attendre de l’IA dépend à la fois de la définition qu’on en retient, de la technologie dont on dispose pour la mettre en œuvre, et des objectifs auxquels on la destine. Il s’ensuit que chacun a de l’IA une vision différente. Ce livre emprunte une voie médiane et aborde l’IA selon autant de points de vue différents qu’il est possible de le faire. Les auteurs prennent leurs distances par rapport à l’emballement frénétique des promoteurs de l’IA, mais sans pour autant souscrire au discours négatif de ses détracteurs. Ils s’attachent à vous présenter la meilleure vision possible de l’IA en tant que technologie. En conséquence, il se peut que vous ayez des attentes quelque peu différentes de ce que vous trouverez dans ce livre, ce qui n’est pas gênant du tout, mais il est essentiel que vous teniez compte de la manière dont cette technologie peut réellement vous servir, plutôt que d’en attendre quelque chose qu’elle ne pourra pas vous apporter.

Intelligence artificielle, qu’est-ce que cela veut dire ?

Avant de pouvoir employer un terme de façon appropriée, il importe de l’avoir bien défini. Quand les gens ne sont pas d’accord sur le sens d’un mot, autant dire que ce mot n’a aucun sens, qu’il n’est rien d’autre qu’une série de caractères. La définition d’une locution (une expression dont la signification n’est pas clairement indiquée par le sens des éléments qui la constituent) est particulièrement importante lorsqu’il s’agit de termes techniques abondamment utilisés dans la presse à différents moments et dans différentes circonstances.



Dire que le sigle IA signifie intelligence artificielle est loin d’être suffisant pour que l’on sache de quoi il est question exactement. Cette expression fait l’objet de nombreuses discussions et de bien des désaccords. Certes, on peut affirmer qu’il s’agit d’un processus artificiel, ne provenant pas d’une source naturelle. Cependant, la référence à l’intelligence est ambiguë, dans le meilleur des cas. Dans les sections qui suivent, les auteurs s’en tiennent à une

définition de l'IA avec laquelle vous ne serez pas nécessairement d'accord, mais c'est en en prenant connaissance que vous pourrez suivre plus facilement le reste du livre.

Discerner l'intelligence

Il existe différentes définitions de l'intelligence. On peut toutefois affirmer que la notion d'intelligence recouvre les activités mentales suivantes :

» **Apprentissage :** être capable d'acquérir et de traiter une nouvelle information.

» **Raisonnement :** être capable de manipuler l'information de différentes manières.

» **Compréhension :** pouvoir apprécier le résultat de la manipulation de l'information.

» **Perception de la vérité :** savoir déterminer la validité de l'information manipulée.

» **Vision des liens :** percevoir la manière dont des données validées interagissent avec d'autres données.

» **Prise en compte du sens :** appliquer des vérités à des situations particulières d'une manière qui soit cohérente avec la relation qui les lie.

» **Distinction entre les faits et les croyances :** pouvoir déterminer si les données sont convenablement étayées par des sources prouvables dont on peut démontrer la validité constante.

Cette liste pourrait facilement être allongée tant et plus, mais elle laisse déjà relativement libre cours à l'interprétation chez quiconque la considère comme viable. Cependant, comme elle le suggère, l'intelligence suit souvent un processus qu'un système informatique peut imiter dans le cadre d'une simulation :

1. **Fixer un objectif en fonction de certains besoins ou de certaines exigences.**
2. **Estimer la valeur d'une information connue à l'appui de l'objectif.**
3. **Obtenir une information complémentaire pouvant appuyer l'objectif.**
4. **Manipuler les données pour leur donner une forme cohérente avec l'information existante.**
5. **Définir les relations et les valeurs de vérité entre les informations existantes et les informations nouvelles.**
6. **Déterminer si l'objectif est atteint.**
7. **Modifier l'objectif à la lumière des nouvelles données et de leur effet sur la probabilité de réussite.**
8. **Répéter les étapes 2 à 7 jusqu'à ce que l'objectif soit atteint (trouvé vrai) ou que les possibilités de l'atteindre soient épuisées (trouvé faux).**



Même si l'on peut créer des algorithmes et assurer l'accès aux données à l'appui de ce processus dans un système informatique, la capacité de l'ordinateur à atteindre le stade de l'intelligence est sévèrement limitée. Ainsi, par exemple, l'ordinateur est incapable de comprendre quoi que ce soit, car il utilise des opérations purement mathématiques pour manipuler les données, de façon strictement mécanique. De même, il ne peut pas facilement distinguer la vérité de son contraire (voir Chapitre 2). En réalité, un ordinateur ne peut mettre pleinement en œuvre aucune des activités mentales que comporte cette liste décrivant ce qu'est l'intelligence.

Pour déterminer ce que recouvre vraiment la notion d'intelligence, il est utile également de catégoriser l'intelligence. Pour exécuter des tâches, l'être humain utilise non pas un seul type, mais plusieurs types d'intelligence. Howard Gardner, de Harvard, en a proposé une liste (voir http://www.pz.harvard.edu/projects/multiple-intelligences). Il est utile de

les connaître pour les mettre en correspondance avec les types de tâches qu'un ordinateur peut exécuter lorsqu'il s'agit de simuler l'intelligence (pour une version modifiée de ces types d'intelligence et pour une description supplémentaire, voir Tableau 1.1).

Tableau 1.1 : Discerner les différents types d'intelligence.

Type	Potentiel de simulation	Outils humains	Description
Visuelle-spatiale	Modéré	Modèles, graphiques, diagrammes, photos, schémas, modélisation en 3D, vidéos, télévision et multimédia	Intelligence de l'environnement physique utilisée par l (entre autres). Pour pouvoir se déplacer, l'être humain environnement physique : ses dimensions et ses cara nécessaire à tout robot et à tout ordinateur portable, m difficile à simuler (comme dans le cas des automobiles précise (comme dans le cas des aspirateurs robots qu rencontrer un obstacle que de disposer d'une forme d direction).
Corporelle-kinesthésique	Modéré à élevé	Matériel spécialisé et objets réels	Mouvements corporels, comme ceux d'un chirurgien d de la précision et une conscience corporelle. Les robo type d'intelligence pour exécuter des tâches répétitive grande précision que les humains, mais parfois avec m de faire la différence entre la capacité humaine, sous équipement chirurgical apportant au praticien une cap un mouvement véritablement indépendant. Dans le pr la démonstration d'une aptitude mathématique dans la dépend des gestes du chirurgien.
Créative	Aucun	Production artistique, nouveaux schémas de pensée, inventions, nouveaux types de composition musicale	La créativité et le fait de développer un nouveau sché une production unique sous forme d'œuvre d'art, de m de produit véritablement nouveau est le résultat de la des schémas de pensée existants, et même les comb sembler être une présentation unique, mais qui ne ser d'un schéma existant produite selon un procédé math une IA aurait besoin de posséder la conscience d'elle- intelligence intrapersonnelle.
Interpersonnelle	Faible à modéré	Téléphone, audioconférence, vidéoconférence, écriture, téléconférence, courrier électronique	Les interactions avec autrui se réalisent à plusieurs ni forme d'intelligence est d'obtenir, d'échanger, de trans informations en fonction de l'expérience des autres. S à des questions simples, c'est grâce à la saisie de mo comprend la question. L'intelligence intervient pour ob localiser les mots-clés pertinents et pour produire ens de ces mots-clés. Croiser les termes dans un tableau fournies par le tableau suppose une intelligence logiq interpersonnelle.
Intrapersonnelle	Faible à modéré	Livres, supports créatifs, journaux, intimité et temps	La capacité de regarder en soi pour comprendre où s fixer des objectifs en fonction de ces intérêts est jusq d'intelligence propre à l'être humain. En tant que mac désirs, d'intérêts, d'exigences ni de facultés créatives. numériques entrantes au moyen d'un ensemble d'algo elle n'a aucune conscience ni aucune compréhension
Linguistique	Faible	Jeux, multimédia,	Le travail avec les mots est un outil essentiel pour la d'informations parlées et écrites est bien plus rapide q

		livres, enregistreurs vocaux et mots parlés	forme de communication. Cette forme d'intelligence re intrants parlés et écrits, le traitement des intrants pour production d'un extrant sous forme de réponse compr cas, l'ordinateur peut à peine analyser l'intrant en term comprend pas la requête, et les réponses qu'il produit compréhensibles du tout. Chez les humains, l'intellige écrite procède de différentes régions du cerveau https://www.sciencedaily.com/releases/20 ce qui signifie que même chez les humains, celui qui f linguistique écrite élevée ne dispose pas nécessairem linguistique orale de niveau similaire. Actuellement, l'o distinguer les facultés linguistiques à l'écrit et à l'oral.
Logique-mathématique	Élevé	Jeux de logique, enquêtes, mystères et casse-tête	Calculer un résultat, effectuer des comparaisons, étud relations, ce sont là des domaines dans lesquels les o Quand un ordinateur bat un humain dans un jeu, c'est que nous voyons, bien qu'il en existe sept. Certes, on d'autres formes d'intelligence, mais on y fait moins att domaine une comparaison entre l'intelligence d'un hu n'est pas une bonne idée.

Découvrir quatre manières de définir l'IA

Comme l'explique la section précédente, la première idée qu'il est important de comprendre est que l'IA, en réalité, n'a rien à voir avec l'intelligence humaine. Certes, certaines formes d'IA sont conçues de manière à simuler l'intelligence humaine, mais il ne s'agit de rien de plus que cela : de la simulation. En matière d'IA, il convient de remarquer l'interaction entre l'activité orientée vers un objectif, le traitement de données qui sert à atteindre cet objectif, et l'acquisition de données qui permet de mieux comprendre l'objectif. L'IA s'appuie sur des algorithmes pour parvenir à un résultat qui peut avoir ou non un lien avec les objectifs des humains et les méthodes permettant d'atteindre ces objectifs. Dans cette optique, on peut catégoriser l'IA de quatre manières :

» **Agir comme des humains :** quand un ordinateur agit comme un être humain, il réussit au mieux au test de Turing, il est alors impossible de faire la différence entre l'ordinateur et l'être humain (voir http://www.turing.org.uk/scrapbook/test.html). Cette catégorie reflète aussi l'image que les médias vous donnent de l'IA. Elle est utilisée pour des technologies comme le traitement automatique du langage naturel, la représentation des connaissances, le raisonnement automatisé et l'apprentissage machine (qui doivent être présentes toutes les quatre pour que le test soit réussi).

» Initialement, le test de Turing devait se dérouler en l'absence de contact physique. Le test total de Turing, qui est nouveau, ne fait pas intervenir de contact physique sous forme d'interrogation de la faculté perceptuelle, ce qui signifie que pour réussir le test, le système informatique doit aussi utiliser à la fois la vision par ordinateur et la robotique. Les techniques modernes reposent sur l'idée d'atteindre l'objectif plutôt que d'imiter totalement les humains. Ainsi, par exemple, ce n'est pas en reproduisant minutieusement le vol des oiseaux que les frères Wright ont réussi à créer un avion : les idées inspirées par l'observation des oiseaux ont abouti à l'aérodynamique, et celle-ci a finalement permis aux hommes de voler. Dans cet exemple, l'objectif était de voler. Les oiseaux et les hommes atteignent cet objectif, mais en utilisant des méthodes différentes.

» **Penser comme des humains** : quand l'ordinateur pense comme un humain, il exécute des tâches pour lesquelles l'être humain doit utiliser son intelligence (par opposition avec les procédures mémorisées), comme conduire une voiture. Pour déterminer si un programme pense comme un humain, il faut disposer d'une méthode qui détermine la

manière dont les humains pensent, et qui est définie par la modélisation cognitive. Le modèle repose sur trois techniques :

  - **Introspection** : détection et documentation des techniques utilisées pour atteindre des objectifs en contrôlant ses propres processus mentaux.
  - **Tests psychologiques :** observation du comportement d'une personne et inclusion dans une base de données de comportements similaires d'autres personnes, compte tenu d'un ensemble similaire de circonstances, d'objectifs, de ressources et de conditions environnementales (entre autres choses).
  - **Imagerie cérébrale :** contrôle direct de l'activité cérébrale par divers moyens mécaniques, comme la tomographie axiale assistée par ordinateur (scanographie), la tomographie par émission de positrons (TEP), l'imagerie à résonance magnétique (IRM) ou la magnétoencéphalographie (MEG).

» Après avoir créé un modèle, on peut écrire un programme qui le simule. Compte tenu de l'importante variabilité des processus de pensée chez l'être humain et de la difficulté de les représenter de manière précise dans un programme, les résultats sont expérimentaux, dans le meilleur des cas. Cette catégorie, penser comme des êtres humains, est souvent utilisée en psychologie et dans d'autres disciplines dans lesquelles la modélisation du processus de pensée humain pour créer des simulations réalistes est essentielle.

» **Penser rationnellement :** l'étude de la façon dont l'être humain pense en s'appuyant sur une certaine norme permet la création de lignes directrices qui décrivent des comportements humains types. Un individu est considéré comme rationnel quand il suit ces comportements en ne s'en écartant que faiblement. Un ordinateur qui pense rationnellement se base sur des comportements enregistrés pour créer un guide décrivant comment interagir avec un environnement basé sur les données dont on dispose. L'objectif de cette approche est de résoudre des problèmes logiquement, dans la mesure du possible. Dans bien des cas, cette approche permettrait la création d'une technique de référence pour la résolution d'un problème, laquelle technique serait ensuite modifiée pour pouvoir résoudre le problème. En d'autres termes, la résolution théorique d'un problème diffère souvent de sa résolution pratique, mais il faut tout de même un point de départ.

» **Agir rationnellement :** étudier la manière dont les humains agissent dans des situations données, confrontés à des contraintes particulières, permet de déterminer quelles techniques seront à la fois efficientes et efficaces. Un ordinateur qui agit rationnellement se base sur les actions enregistrées pour interagir avec un environnement en fonction des conditions, des facteurs environnementaux et des données existantes. Les actes rationnels, comme les pensées rationnelles, dépendent d'une solution de principe qui peut ne pas se révéler utile en pratique. Cependant, les actes rationnels constituent effectivement une base de référence à partir de laquelle l'ordinateur peut commencer à accomplir la réalisation d'un objectif.

PROCESSUS HUMAINS ET PROCESSUS RATIONNELS

Les processus humains diffèrent des processus rationnels en termes de résultat. Un processus est rationnel s'il accomplit toujours ce qu'il convient d'accomplir en fonction de l'information courante, compte tenu d'une mesure de performance idéale. En résumé, les processus rationnels appliquent le texte à la lettre en considérant que le texte est correct. Les processus humains font appel à l'instinct, à l'intuition et à d'autres variables qui ne reflètent pas nécessairement le livre et peuvent même ne pas tenir compte du tout des données existantes. À titre d'exemple, la façon rationnelle de conduire une voiture consiste à toujours respecter le code de la route. Or, le trafic n'est pas rationnel. En appliquant les textes à la lettre, on finit par se retrouver coincé quelque part, car les autres automobilistes ne respectent pas tous les règles de façon scrupuleuse. Pour pouvoir fonctionner, la voiture sans conducteur doit donc se comporter comme les humains plutôt que de façon rationnelle.

À travers les catégories utilisées pour définir l'IA, on peut étudier différentes manières de l'exploiter. Cependant, les systèmes de classification de l'IA sont parfois arbitraires et insatisfaisants. Ainsi, par exemple, certains considèrent que l'IA est soit forte (intelligence générale, ou capacité de s'adapter à une variété de situations), soit faible (intelligence spécifique, orientée vers la bonne exécution d'une tâche particulière). Le problème d'une IA forte est qu'elle n'exécute aucune tâche de façon performante, tandis qu'une IA faible est trop spécifique pour pouvoir accomplir des tâches de façon indépendante. Quoi qu'il en soit, distinguer seulement deux types ne fait pas l'affaire, même d'un point de vue général. Les quatre types de classification proposés par Arend Hintze (voir `http://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-toself-aware-beings-67616`) constituent une meilleure base pour la compréhension de l'IA :

- **Les machines réactives** : les machines qui battent les humains aux échecs et celles utilisées dans les jeux télévisés sont des exemples de machines réactives. Une machine réactive n'a pas de mémoire ni d'expérience sur laquelle fonder une décision. Elle n'utilise que la puissance de calcul et des algorithmes intelligents pour recréer chaque décision à chaque fois. C'est un exemple d'IA faible utilisée dans un but précis.
- **La mémoire limitée :** une voiture sans conducteur ou un robot autonome ne peuvent pas prendre le temps d'aboutir à chaque décision à partir de rien. Ces machines exploitent une petite quantité de mémoire pour mettre en œuvre la connaissance expérientielle de diverses situations. Quand la machine reconnaît une situation à laquelle elle a déjà fait face, elle peut utiliser son expérience pour réduire son temps de réaction et mobiliser davantage de ressources, afin de prendre de nouvelles décisions qui n'ont pas encore été élaborées. C'est un exemple du niveau actuel de l'IA forte.
- **La théorie de l'esprit :** le type de compréhension caractérisant une machine qui peut évaluer à la fois les objectifs qu'elle travaille à atteindre et les objectifs potentiels d'autres entités au sein du même environnement est envisageable aujourd'hui dans une certaine mesure, mais pas sous une forme commerciale. Cependant, pour que les automobiles sans conducteur puissent devenir véritablement autonomes, il faut que ce niveau d'IA soit entièrement développé. Une voiture sans conducteur aurait besoin non seulement de savoir qu'elle doit aller d'un point donné à un autre, mais aussi, de connaître intuitivement les objectifs potentiellement conflictuels des conducteurs des véhicules qui gravitent autour d'elle afin de pouvoir réagir en conséquence.
- **La conscience de soi :** c'est le type d'IA qu'on voit dans les films. Or, une telle IA repose sur des technologies qui ne sont même pas envisageables dans l'avenir, sachant qu'une telle machine devrait posséder à la fois un sens de la perception de soi et une conscience. En outre, au lieu de simplement supposer les objectifs des autres entités en

> fonction de leurs réactions et de l'environnement, ce type de machine serait capable de déduire les intentions des autres d'après une connaissance expérientielle.

Comprendre l'histoire de l'IA

Les sections précédentes de ce chapitre vous donnent un aperçu de l'intelligence du point de vue de l'humain et vous expliquent que l'informatique actuelle est lamentablement inadéquate pour ce qui est de simuler cette intelligence, et qu'elle peut encore moins devenir elle-même intelligente. Or, le désir de créer des machines intelligentes (ou, dans l'Antiquité, des idoles) existe depuis que l'homme existe. C'est le désir de ne pas être seul dans l'univers, d'avoir avec qui communiquer sans subir les incohérences des autres humains, et ce désir est profond. Naturellement, un seul livre ne saurait embrasser l'ensemble de l'histoire humaine. Les sections qui suivent présentent un bref aperçu pertinent de l'histoire des efforts actuels en matière d'IA.

Tout a commencé à Dartmouth avec la logique symbolique

Les premiers ordinateurs n'étaient que des ordinateurs au sens propre du terme, c'est-à-dire des calculateurs. Ils imitaient la faculté humaine de manipuler des symboles pour exécuter des opérations mathématiques élémentaires comme l'addition. Par la suite, les machines sont devenues capables de raisonnement logique au moyen de comparaisons (consistant à déterminer si une valeur est plus grande qu'une autre). Cependant, il revenait aux humains de définir l'algorithme utilisé pour exécuter les calculs, de fournir à la machine les données requises dans le format adéquat, puis d'interpréter le résultat. Durant l'été 1956, des scientifiques ont participé à un atelier de travail sur le campus de Dartmouth College en vue d'aller plus loin. Ils ont prédit qu'en l'espace d'à peine une génération, les machines allaient pouvoir raisonner aussi bien qu'on pourrait le souhaiter. Ils se sont trompés. Ce n'est qu'aujourd'hui que nous savons fabriquer des machines capables de raisonnements mathématiques et logiques comparables aux nôtres (il a fallu que les ordinateurs acquièrent au moins six autres formes d'intelligence avant qu'il leur soit possible d'imiter un tant soit peu l'intelligence humaine).

Le problème que n'avaient pas résolu les scientifiques de cette époque, à Dartmouth College ou ailleurs, était lié à la capacité matérielle, c'est-à-dire à la puissance et à la rapidité de calcul nécessaires pour réaliser une simulation. En réalité, ce n'est pas là tout le problème. C'était aussi une question de matériel, certes, mais il n'est pas possible de simuler des processus que l'on ne comprend pas. Quoi qu'il en soit, la raison pour laquelle l'IA est relativement efficace aujourd'hui est que le matériel est finalement devenu assez performant pour pouvoir effectuer le nombre de calculs requis.



Dans ces premières tentatives, le plus gros problème (qui reste un problème considérable aujourd'hui) était le fait que nous ne comprenions pas assez bien comment l'être humain raisonne pour pouvoir simuler cette activité d'une quelconque manière – en supposant qu'une simulation soit possible. Revenons sur les problèmes évoqués précédemment dans ce chapitre concernant le vol habité. Plutôt que d'imiter les oiseaux, les frères Wright ont cherché à comprendre les processus qui leur permettaient de voler. C'est ainsi qu'ils ont donné naissance à l'étude de l'aérodynamique. Par conséquent, quand on entend dire que la prochaine innovation majeure en IA est imminente, alors même qu'il n'existe aucune description des processus en jeu, la réalité est que l'innovation en question n'est pas imminente du tout.

Pour continuer avec les systèmes experts

Les systèmes experts ont fait leur apparition dans les années 1970, puis à nouveau dans les années 1980 dans le cadre d'une tentative de réduire, grâce à la connaissance d'experts, les exigences de calculs posées par l'IA. Cela a donné lieu à un certain nombre de représentations de systèmes experts, qui pouvaient être à base de règles (avec des décisions fondées sur des déclarations de type si… alors) ou de cadres (avec des bases de données organisées en structures hiérarchiques d'informations génériques) ou fondés sur la logique (établissement de relations en fonction d'une théorie établie). L'avènement des systèmes experts est une étape importante, car il s'agit des premières mises en œuvre vraiment utiles et probantes de l'IA.



Aujourd'hui, les systèmes experts sont encore utilisés (même si on ne les désigne plus par ce terme). La fonction de vérificateur d'orthographe et de correcteur grammatical de votre logiciel de traitement de texte, par exemple, est un système expert. Le correcteur grammatical, en particulier, est essentiellement basé sur l'application de règles. Il serait intéressant de chercher quelles sont les autres utilisations pratiques actuelles des systèmes experts, dans des applications quotidiennes.

Un problème que posent les systèmes experts est qu'ils peuvent être difficiles à mettre au point et à entretenir. Leurs premiers utilisateurs étaient obligés de maîtriser des langages de programmation spécialisés comme List Processing (Lisp) ou Prolog. Certaines entreprises du secteur ont pu mettre des systèmes experts entre les mains de programmeurs novices ou moins expérimentés grâce à l'utilisation de produits qui étaient conçus selon une approche fondée sur les règles, comme VP-Expert https://www.amazon.com/exec/obidos/ASIN/155622057X/datacservip0f-20/). Cependant, les fonctionnalités offertes par ces produits qui utilisaient des bases de connaissance réduites étaient généralement très limitées.

Au cours des années quatre-vingt-dix, l'expression *système expert* a cessé peu à peu d'être utilisée. Les systèmes experts commençaient à être considérés comme un échec, bien qu'en réalité leur succès soit tel qu'ils étaient désormais intégrés dans les applications pour lesquelles ils avaient été conçus afin de les compléter. Si l'on prend l'exemple du traitement de texte, il fut un temps où il était nécessaire de se procurer séparément un logiciel de vérification grammaticale comme RightWriter (http://www.right-writer.com/). Ce genre d'application s'est révélée si utile (même si les résultats ne sont pas toujours pertinents), que les logiciels de traitement de texte intègrent maintenant un vérificateur de grammaire (voir https://www.presse-citron.net/google-annonce-un-correcteur-grammatical-base-sur-lintelligence-artificielle/).

Passer le cap des hivers de l'IA

L'expression « hiver de l'IA » désigne une période au cours de laquelle les financements alloués au développement de l'IA se sont réduits. De façon générale, les promoteurs de l'IA ont eu tendance à exagérer les possibilités qu'elle offrait, et ainsi, à inciter des gens qui n'avaient aucune connaissance technique mais qui disposaient de fonds importants à investir. Par la suite, l'IA ne répondait pas aux attentes, si bien que le bien-fondé de son financement était remis en cause. Au cours du temps, ce cycle s'est répété, avec à chaque fois des conséquences funestes en termes de progrès véritable.

Si l'IA connaît actuellement un nouvel engouement, c'est grâce à l'apprentissage machine, une technologie qui permet aux ordinateurs de tirer des leçons des données. Quand un ordinateur est capable d'apprendre à partir des données qu'il traite, cela signifie qu'au lieu de dépendre d'un programmeur humain pour définir et effectuer des opérations (des tâches), il les déduit directement d'exemples qui lui indiquent comment il doit se comporter. Tout se passe comme lorsqu'un bébé prend exemple sur le comportement qu'on lui montre. L'apprentissage machine n'est certes pas exempt de pièges, sachant que l'ordinateur peut apprendre à faire des choses incorrectes si l'enseignement qui lui est dispensé est bâclé.

Cinq tribus de scientifiques travaillent actuellement sur des algorithmes d'apprentissage machine, chacune selon un angle différent (voir la section « Éviter l'emphase concernant l'IA », plus loin dans ce chapitre). En ce moment, la solution la plus fructueuse est l'*apprentissage profond*, une technologie qui vise à imiter le cerveau humain. L'apprentissage profond est rendu possible par l'existence d'ordinateurs puissants, d'algorithmes plus intelligents et de grandes séries de données produites par la numérisation de notre société, et par les énormes investissements réalisés par des entreprises comme Google, Facebook ou Amazon, qui savent tirer parti de cette renaissance de l'IA.

On entend dire que l'hiver de l'IA est fini grâce à l'apprentissage profond, et c'est vrai pour l'instant. Cependant, compte tenu de la vision que les gens ont généralement de l'IA, on peut facilement anticiper une nouvelle vague de critiques, à moins que ses promoteurs mettent un frein à leur rhétorique. L'IA peut faire des choses remarquables, mais plutôt dans des applications à caractère routinier, comme l'explique la section suivante.

Recenser les applications de l'IA

L'IA est aujourd'hui utilisée dans un très grand nombre d'applications. Le seul problème, c'est que la technologie est si performante qu'on ne sait même pas qu'elle existe. Vous serez peut-être surpris de découvrir qu'un certain nombre d'appareils que vous possédez l'utilisent déjà. Certains thermostats intelligents, par exemple, créent automatiquement des programmes pour vous en fonction de la manière dont vous contrôlez manuellement la température. De même, les appareils que vous pouvez commander par la voix apprennent à reconnaître votre façon de parler, afin de mieux interagir avec vous. L'IA est maintenant présente dans votre voiture, et surtout sur votre lieu de travail. En réalité, l'IA trouve des millions d'utilisations, mais qui passent toutes inaperçues même quand elles sont pour ainsi dire spectaculaires. En voici seulement quelques exemples :

» **La détection des fraudes** : peut-être vous est-il arrivé de recevoir un message de l'entreprise qui gère votre carte bancaire, vous demandant si c'était bien vous qui veniez de procéder à tel ou tel achat. Ce n'était pas de la curiosité de sa part : l'IA intégrée à la gestion du code de votre prestataire de services bancaires avait détecté un type de dépense inhabituel et vous alertait, au cas où votre carte aurait été utilisée par quelqu'un d'autre.

» **La planification des ressources :** nombreuses sont les organisations qui doivent programmer une utilisation efficiente de leurs ressources. Un hôpital, par exemple, peut avoir besoin de déterminer où un patient doit être installé, en tenant compte des besoins de celui-ci, de la disponibilité des spécialistes qualifiés et de la durée d'hospitalisation prévue par le médecin.

» **L'analyse complexe :** souvent, le recours à une analyse complexe s'impose en raison du nombre très élevé de facteurs à prendre en compte. Ainsi, par exemple, un même ensemble de symptômes n'indique pas nécessairement un seul problème. Le médecin peut avoir besoin de s'appuyer sur une analyse approfondie pour être en mesure de formuler le bon diagnostic dans un délai convenable et sauver la vie du patient.

» **L'automatisation** : dans toute forme d'automatisation, faire intervenir l'IA peut être souhaitable pour pouvoir faire face à des changements ou à des événements imprévus. Un problème qui se pose parfois aujourd'hui avec certains types d'automatisation est qu'un événement imprévu, par exemple un objet qui n'est plus à la bonne place, peut entraîner l'arrêt du processus automatisé. L'IA peut permettre à l'automatisation de gérer les imprévus et d'assurer la continuité du processus, comme s'il ne s'était rien passé.

» **Le service à la clientèle :** aujourd'hui, quand vous appelez le service clients d'une entreprise, il est même possible qu'aucun être humain n'intervienne à l'autre extrémité de la ligne. Les systèmes automatisés sont capables d'exécuter des scripts et d'utiliser

diverses ressources pour donner suite valablement à une grande majorité des questions. Grâce à des inflexions de voix appropriées (produites également par l'IA), il se peut que vous ne puissiez même pas vous rendre compte que c'est à un ordinateur que vous parlez.

» **Les systèmes de sécurité :** aujourd'hui, des systèmes de sécurité installés sur diverses sortes de machines utilisent l'IA, par exemple pour prendre le contrôle d'un véhicule en cas de nécessité. Les systèmes de freinage automatique, notamment, utilisent l'IA pour arrêter le véhicule en fonction de toutes les données que ce véhicule peut fournir, comme la direction d'un dérapage.

» **L'efficience des machines :** l'IA peut permettre de piloter une machine de manière à obtenir une efficacité maximum. Elle contrôle l'utilisation des ressources de telle manière que le système ne dépasse pas une certaine vitesse, par exemple. La quantité d'énergie nécessaire est utilisée avec toute la précision voulue pour fournir les services demandés.

Éviter l'emphase concernant l'IA

Dans ce chapitre, il est beaucoup question du battage fait autour de l'IA. Et cependant, ce chapitre ne peut qu'en donner une faible idée. Si vous regardez des films comme *Her* (https://www.amazon.com/exec/obidos/ASIN/B00H9HZGQ0/datacservip0f-20/) ou *Ex Machina* (https://www.amazon.com/exec/obidos/ASIN/B00XI057M0/datacservip0f-20/), vous serez incité à surestimer les avancées en matière d'IA. Le problème est qu'en réalité, l'IA n'en est qu'à ses débuts, et que toutes les applications que montrent ces films ne sont que le produit d'une imagination débridée.

Peut-être avez-vous entendu parler de ce qu'on appelle la singularité, une hypothèse dont découle tout ce qui peut être présenté comme réalisable dans les médias et dans les films. La *singularité* est essentiellement un algorithme maître qui englobe les cinq « tribus » de l'apprentissage machine. Pour réaliser ce que suggèrent ces sources, la machine doit être capable d'apprendre comme le ferait un être humain, en utilisant les sept types d'intelligence dont il est question dans la section « Discerner l'intelligence », précédemment dans ce chapitre.

Voici les cinq tribus de l'apprentissage machine :

» **Les symbolistes** : cette tribu s'inspire de la logique et de la philosophie, et résout les problèmes par déduction inverse.

» **Les connexionnistes :** cette tribu trouve son origine dans les neurosciences et résout les problèmes par rétropropagation du gradient.

» **Les évolutionnaires** : la tribu des évolutionnaires s'inspire de la biologie évolutive et s'appuie sur la programmation génétique pour résoudre les problèmes.

» **Les bayésiens** : cette tribu trouve son origine dans les statistiques et résout les problèmes à l'aide de l'inférence probabiliste (ou bayésienne).

» **Les analogistes** : cette tribu procède de la psychologie et utilise les machines à vecteurs de support (SVM) pour résoudre les problèmes.

Le but ultime de l'apprentissage machine est d'associer les technologies et les stratégies que recouvrent les cinq tribus, de manière à créer un unique algorithme (l'algorithme maître) qui soit capable d'apprendre quelque chose. Naturellement, nous sommes loin d'avoir atteint ce but. Cependant, des scientifiques comme Pedro Domingos (http://homes.cs.washington.edu/~pedrod/) y travaillent en ce moment.

Pour compliquer encore les choses, il n’est pas dit que les cinq tribus soient capables de produire suffisamment d’informations pour pouvoir résoudre le problème de l’intelligence humaine. Par conséquent, la création d’algorithmes maîtres pour les cinq tribus ne débouchera peut-être pas encore sur la singularité. Ici, vous pourriez être surpris de constater à quel point les gens ignorent la manière dont ils pensent et la raison pour laquelle ils pensent d’une certaine manière. Toutes les rumeurs qui peuvent vous parvenir concernant la possibilité que l’IA gouverne le monde, ou devienne supérieure aux humains, sont absolument dénuées de fondement.

Relier l’IA au système informatique

Pour voir comment fonctionne l’IA, il vous faut un système informatique, une application comportant le logiciel requis, et une base de connaissances. N’importe quel ordinateur ou système est susceptible de faire l’affaire, pourvu qu’il soit doté d’une puce : pour certaines applications, un téléphone intelligent convient tout aussi bien qu’un ordinateur de bureau. Bien sûr, une entreprise comme Amazon ne pourra pas utiliser un téléphone pour conseiller ses clients dans leurs décisions d’achat : pour cela, elle aura besoin d’un système informatique de grande dimension.

La taille du système informatique est directement proportionnelle à la quantité de travail que l’IA est censée accomplir. L’application peut aussi varier par sa taille, sa complexité et même sa localisation. Par exemple, si vous êtes une entreprise et si vous voulez analyser les données relatives aux clients en vue de déterminer le meilleur argumentaire de vente, vous utiliserez probablement pour cela une application sur serveur. Si vous êtes un client, et si vous cherchez sur Amazon des produits qui compléteront votre commande en cours, l’application ne résidera même pas sur votre ordinateur : vous y accéderez par le biais d’une application située sur les serveurs d’Amazon.

La base de connaissances varie aussi en termes de localisation et de taille. Plus les données sont complexes, plus vous pouvez en tirer d’informations, mais plus vous avez besoin de les manipuler. Dans la gestion des connaissances, rien n’est gratuit. Le compromis entre la localisation et le temps est important aussi. Une connexion réseau vous ouvre l’accès à une vaste base de connaissances en ligne, mais cela vous demande du temps en raison de la latence des connexions. D’un autre côté, bien souvent une base de données locale, si elle permet des accès rapides, contient moins de données détaillées.

Chapitre 2

Définir le rôle des données

DANS CE CHAPITRE…

- » **Considérer les données comme une ressource universelle**
- » **Obtenir et manipuler des données**
- » **Rechercher des erreurs dans les données**
- » **Définir les limites de l'acquisition de données**

Il n'y a pas grand-chose de nouveau concernant les données. Toute application intéressante conçue pour être utilisée sur un ordinateur fonctionne avec des données qui lui sont associées. Les données peuvent se présenter sous des formes variées. Elles peuvent être organisées ou non. Ce qui a changé, ce sont les volumes de données. Nous avons maintenant accès, et cela peut faire peur, à des quantités considérables de données sur pratiquement tous les aspects de la vie des gens, parfois jusqu'à un niveau de détail que les gens ne réalisent pas. Par ailleurs, l'utilisation de matériels sophistiqués et les progrès réalisés dans les algorithmes font que les données sont aujourd'hui la ressource universelle de l'IA.

Pour pouvoir travailler avec des données, il faut tout d'abord les obtenir. Aujourd'hui, les applications recueillent les données manuellement, comme par le passé, ou automatiquement, à l'aide de nouvelles méthodes. Cependant, il n'existe pas simplement une ou deux techniques de recueil de données : les méthodes

s’échelonnent dans un continuum, depuis la collecte purement manuelle jusqu’à la collecte entièrement automatique.

Les données brutes ne conviennent généralement pas aux fins d’analyse. Ce chapitre vous permet aussi de comprendre pourquoi il est nécessaire de manipuler les données et de les mettre en forme de manière à ce qu’elles répondent à certaines exigences spécifiques. Vous allez aussi découvrir le besoin de définir la valeur de vérité des données de telle sorte que les résultats de l’analyse correspondent aux objectifs fixés initialement pour les applications.

De façon singulière, il importe également de prendre en compte des limites à l’acquisition de données. Il n’existe encore aucune technologie qui permette de lire les pensées de quelqu’un par voie télépathique. Bien sûr, il existe aussi d’autres limites. Vous les connaissez probablement déjà, la plupart du temps, mais vous n’y avez peut-être pas songé.

Constater que les données sont aujourd’hui omniprésentes

La révolution des « grandes données » (le « big data ») n’est pas un simple concept à la mode utilisé pour vendre de nouveaux moyens de stocker et d’analyser les données, c’est une réalité de tous les jours et une force conductrice de notre temps. Vous avez pu entendre parler des grandes données en lisant des publications scientifiques ou économiques, et vous vous êtes peut-être demandé ce que recouvrait ce terme. D’un point de vue technique, le *big data* fait référence à des ensembles de données informatiques qui sont si vastes et si complexes que les applications informatiques ne peuvent pas les

traiter en utilisant des capacités de stockage supplémentaires ni une plus grande puissance de traitement.

Les grandes données rendent nécessaire une révolution dans les systèmes de stockage et de manipulation des données. Il s'agit de ce que l'on peut faire avec des données sur un plan plus qualitatif (non seulement faire plus, mais faire mieux). Du point de vue humain, les systèmes informatiques stockent les grandes données sous différents formats, mais pour l'ordinateur, les données restent un flux de un et de zéros (c'est le langage de base des ordinateurs). On peut surtout distinguer deux types de données, selon la manière dont elles sont produites et utilisées. Il y a celles qui sont organisées selon une structure évidente (on sait exactement ce qu'elles contiennent et où trouver chaque élément de donnée) et celles qui ne sont pas structurées (on a une idée de ce qu'elles contiennent, mais sans savoir exactement comment elles sont organisées).

Les bases de données sont des exemples types de données structurées. L'information apparaît sous forme de colonnes et chaque colonne contient un type spécifique d'information. Les données sont souvent structurées selon la conception du système de collecte. Elles sont recueillies de manière sélective, et enregistrées là où il convient. Par exemple, le comptage du nombre de personnes qui achètent un certain produit sera inscrit dans une colonne particulière d'un tableau particulier, dans une base de données particulière. De la même manière que dans une bibliothèque, si vous savez de quelles données vous avez besoin, vous pouvez les trouver tout de suite.

Les données non structurées sont des images, des vidéos et des enregistrements sonores. Vous pouvez utiliser une forme non structurée pour un texte auquel vous attacherez des caractéristiques comme sa taille, sa date de création et son type de contenu. Généralement, on ignore où

exactement les données apparaissent dans un ensemble de données non structuré, car elles apparaissent sous forme de successions de un et de zéros qu'une application doit interpréter ou visualiser.



Transformer des données non structurées en données structurées peut coûter beaucoup de temps et d'efforts, et nécessiter le travail de nombreuses personnes. Les grandes données sont le plus souvent non structurées et stockées telles quelles, sauf lorsqu'on les transforme en données structurées.

Cet amas de données abondantes et complexes n'est pas apparu du jour au lendemain. Il a fallu du temps pour développer la technologie permettant de stocker de telles quantités de données. En outre, il a fallu du temps pour diffuser la technologie qui produit les données, c'est-à-dire les ordinateurs, les capteurs, les téléphones mobiles intelligents, l'Internet et les services du *World Wide Web*. Les sections qui suivent vous permettent de comprendre pourquoi les données sont devenues aujourd'hui une ressource universelle.

Comprendre les implications de la loi de Moore

En 1965, dans un article du magazine *Electronics* intitulé « Cramming More Components Onto Integrated Circuits » (http://ieeexplore.ieee.org/document/4785860/), Gordon Moore, cofondateur d'Intel et de Fairchild Semiconductor, avait prédit que le nombre de composants dans les circuits intégrés allait doubler tous les ans au cours de la décennie à venir. À l'époque, l'électronique était dominée par les transistors. Pouvoir fourrer davantage de transistors dans un circuit intégré (CI), c'était pouvoir accroître la performance et l'utilité des

appareils électroniques. Cette évolution, appelée l'intégration, supposait un processus de miniaturisation poussé (la miniaturisation consistant à fabriquer un même circuit de telle manière qu'il soit toujours plus petit). Les ordinateurs d'aujourd'hui ne sont pas beaucoup plus petits que ceux d'il y a dix ans, mais ils sont considérablement plus puissants. Il en est de même des téléphones mobiles. Même s'ils ont les mêmes dimensions que les modèles précédents, ils peuvent maintenant exécuter davantage de tâches.

Ce que Moore avait annoncé dans cet article est resté vrai pendant longtemps. Dans l'industrie des semi-conducteurs, on a appelé cela la loi de Moore (voir http://www.mooreslaw.org/). Le doublement a bien eu lieu au cours de la première décennie, comme cela avait été prédit. En 1975, Moore corrigea sa déclaration et prédit un doublement tous les deux ans. La Figure 2.1 montre les effets de ce doublement. Cette fréquence du doublement reste vérifiée actuellement, bien qu'il soit communément admis que la tendance observée ne se prolongera pas après 2020. En effet, à partir de 2012, un décalage a commencé à apparaître entre les accélérations attendues et les possibilités de progrès en matière de miniaturisation dans l'industrie des semi-conducteurs.

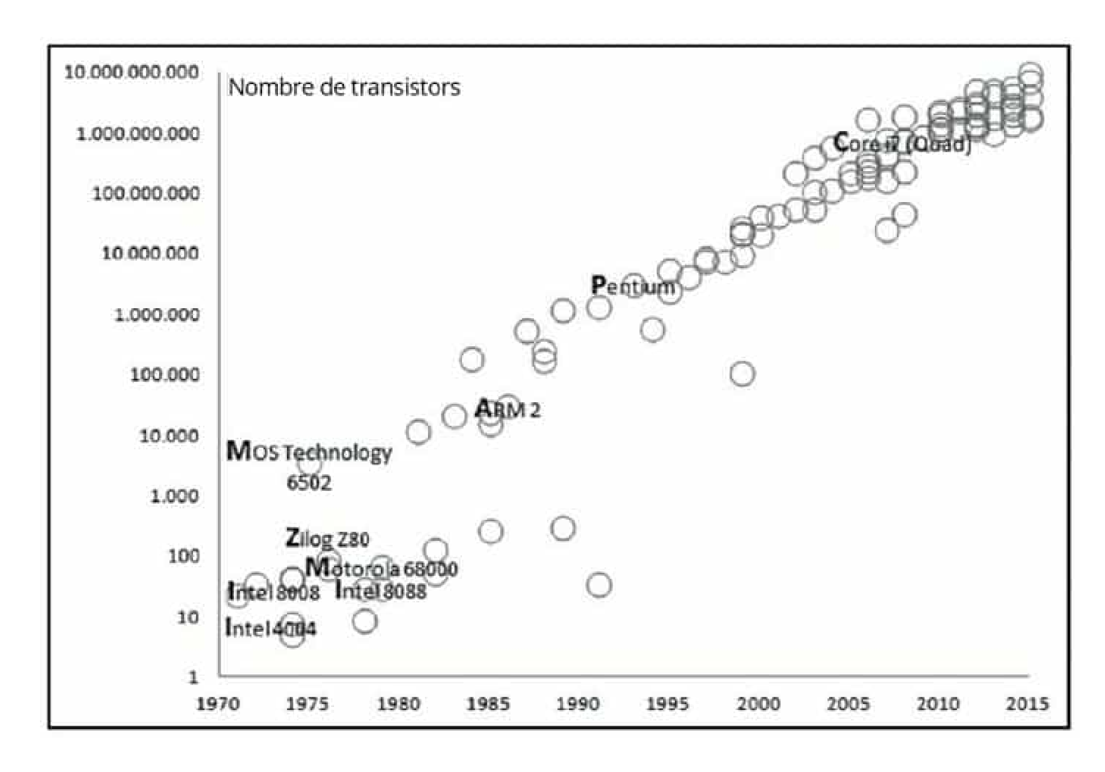


FIGURE 2.1 : De plus en plus de transistors dans un processeur.

Il existe des barrières physiques à l'accroissement du nombre de circuits que l'on peut intégrer dans un CI en utilisant les composants en silicium actuels, sachant qu'on ne peut pas indéfiniment fabriquer toujours plus petit. L'innovation se poursuit cependant (voir http://www.nature.com/news/the-chips-aredown-for-moores-law-1.19338). Dans l'avenir, la loi de Moore pourrait ne plus s'appliquer, si l'industrie adopte une nouvelle technologie (comme la fabrication de composants en utilisant des transistors optiques à base de laser à la place des transistors électroniques : à propos de l'ordinateur optique, lire l'article à l'adresse http://www.extremetech.com/extreme/187746-by-2020-you-could-have-an-exascalespeed-of-light-optical-computeron-your-desk). Ce qu'il convient de retenir, c'est que depuis 1965, le doublement du nombre de composants tous les deux ans s'est accompagné de grandes avancées dans le domaine de l'électronique numérique, avec de profondes

répercussions en termes d'acquisition, de stockage, de manipulation et de gestion de données.

La loi de Moore exerce un impact direct sur les données. Il s'agit tout d'abord d'appareils plus intelligents. Plus les appareils sont intelligents, plus grande est leur diffusion (comme le montre l'omniprésence actuelle de l'électronique). Plus grande est leur diffusion, moins les prix sont élevés, ce qui engendre l'utilisation généralisée de puissantes machines programmables et de petits capteurs selon une boucle sans fin. Compte tenu des quantités considérables de mémoire informatique disponible et de plus grands disques de stockage de données, la disponibilité des données est expansion, qu'il s'agisse de sites Internet, d'enregistrements de transactions, de mesures quantitatives, d'imagerie numérique ou d'autres sortes de données.

Utiliser des données partout

Pour leurs expériences, les scientifiques ont besoin d'ordinateurs plus puissants que la moyenne des utilisateurs. Ils traitaient déjà d'impressionnants volumes de données des années avant que le terme « big data » soit inventé. L'Internet ne produisait pas encore les ensembles considérables de données qu'il produit aujourd'hui. Il ne faut pas oublier que les grandes données ne sont pas une simple mode lancée par les fabricants d'ordinateurs et de logiciels, mais qu'elles jouent un rôle essentiel dans divers domaines scientifiques comme l'astronomie (et les missions spatiales), les satellites (pour la surveillance et le contrôle), la météorologie, la physique (les accélérateurs de particules) et la génomique (les séquences d'ADN).

Si les applications de l'IA peuvent être propres à un domaine scientifique particulier comme Watson d'IBM, qui

s’enorgueillit d’une capacité de diagnostic impressionnante, car cette application est capable d’assimiler des informations provenant de plusieurs millions d’articles scientifiques sur les maladies et les traitements médicaux, les applications actuelles de l’IA présentent souvent des aspects plus ordinaires. Elles sont surtout appréciées pour leur capacité de reconnaître des objets, de parcourir des chemins ou de comprendre ce qu’on leur dit. La contribution des données au nouvel essor de l’IA et à ces orientations ne provient pas des sources classiques de données scientifiques.

Aujourd’hui, l’Internet produit et distribue de nouvelles données en grandes quantités. La production quotidienne de données est actuellement estimée aux alentours de 2,5 quintillions (un nombre comportant 18 zéros) d’octets, la part du lion revenant aux données non structurées, notamment les données audio et vidéo. Toutes ces données sont liées à des activités, des sensations, des expériences et des relations humaines courantes. En parcourant ces données, une IA peut facilement apprendre à raisonner et à agir en imitant l’être humain. Voici quelques exemples particulièrement intéressants :

» De vastes collections de visages et d’expressions provenant de photos et de vidéos publiées sur des sites de réseaux sociaux comme Facebook, YouTube et Google donnent des informations sur le sexe, l’âge, les goûts et éventuellement les préférences sexuelles, les orientations politiques ou le QI (voir https://www.theguardian.com/technology/2017/sep/12/artificial-intelligence-face-recognition-michal-kosinski).

» Les renseignements médicaux et les données biométriques provenant des montres intelligentes, et détenus par des entreprises privées, permettent de

mesurer des indicateurs corporels comme la température et le pouls chez des patients malades ou en bonne santé.

» De vastes bases de données provenant de sources comme les réseaux sociaux et les moteurs de recherche permettent de savoir comment les gens interagissent entre eux et quels sont leurs centres d'intérêt. Ainsi, d'après une étude du Centre de psychométrie de l'université de Cambridge, les interactions sur Facebook contiendraient beaucoup de données sur les relations intimes.

» Des informations sur notre façon de nous exprimer sont enregistrées par les téléphones mobiles. OK Google, par exemple, une fonction qui équipe les téléphones mobiles utilisant le système Android, enregistre systématiquement les questions, et parfois même davantage : https://qz.com/526545/googlesbeen-quietly-recording-your-voice-heres-how-to-listen-to-anddelete-the-archive/.

Le nombre d'appareils que les utilisateurs connectent à l'Internet augmente tous les jours, et à chaque fois, de nouvelles données personnelles sont stockées. Il existe maintenant des assistants personnels qu'on installe chez les particuliers, comme Amazon Echo ou d'autres appareils domestiques intelligents intégrés, qui permettent au consommateur de commander ses appareils ménagers et de se faciliter la vie à domicile. Ce n'est là que la partie émergée de l'iceberg, sachant que beaucoup d'autres outils et objets d'utilisation quotidienne deviennent interconnectés (du réfrigérateur à la brosse à dents) et capables de traiter, d'enregistrer et de transmettre des données. L'Internet des objets est en train de devenir une réalité. Des spécialistes ont estimé qu'il y aurait en 2020 six fois plus d'objets connectés que d'êtres

humains, mais ces chiffres sont déjà remis en question par des chercheurs et des panels d'experts (http://www.gartner.com/newsroom/id/3165317).

Mettre les algorithmes en application

Le genre humain se trouve actuellement confronté à des volumes de données inédits, produits par un matériel informatique de plus en plus petit et de plus en plus puissant. Par ailleurs, de plus en plus, les données sont traitées et analysées par ces mêmes ordinateurs que ce processus a permis de développer et de généraliser. Cela peut sembler évident, mais les données sont devenues omniprésentes à tel point que leur valeur ne réside plus seulement dans l'information qu'elles contiennent (comme dans le cas des données stockées dans la base de données d'une entreprise pour les besoins de ses activités quotidiennes), mais plutôt dans leur utilisation comme moyen de créer une nouvelle valeur : elles sont considérées comme le « nouvel or noir ». Cette nouvelle valeur réside principalement dans la manière dont les applications gèrent, stockent et récupèrent les données, et dont on peut effectivement les utiliser au moyen d'algorithmes intelligents.

En matière de données, les algorithmes et l'IA ont changé les règles du jeu. Comme mentionné au chapitre précédent, les algorithmes de l'IA ont adopté des approches différentes, depuis les algorithmes simples jusqu'au raisonnement symbolique fondé sur la logique et aux systèmes experts. Depuis quelques années, ils sont devenus des réseaux de neurones, et dans leur forme la plus mature, ils ont inauguré l'apprentissage profond. Au cours de cette transition méthodologique, les données ont

cessé d'être simplement de l'information traitée par des algorithmes prédéterminés pour devenir ce qui rend les algorithmes utiles pour une tâche. Au lieu d'être simplement la matière première pour l'élaboration de la solution, elles sont devenues ce qui crée la solution (voir Figure 2.2).

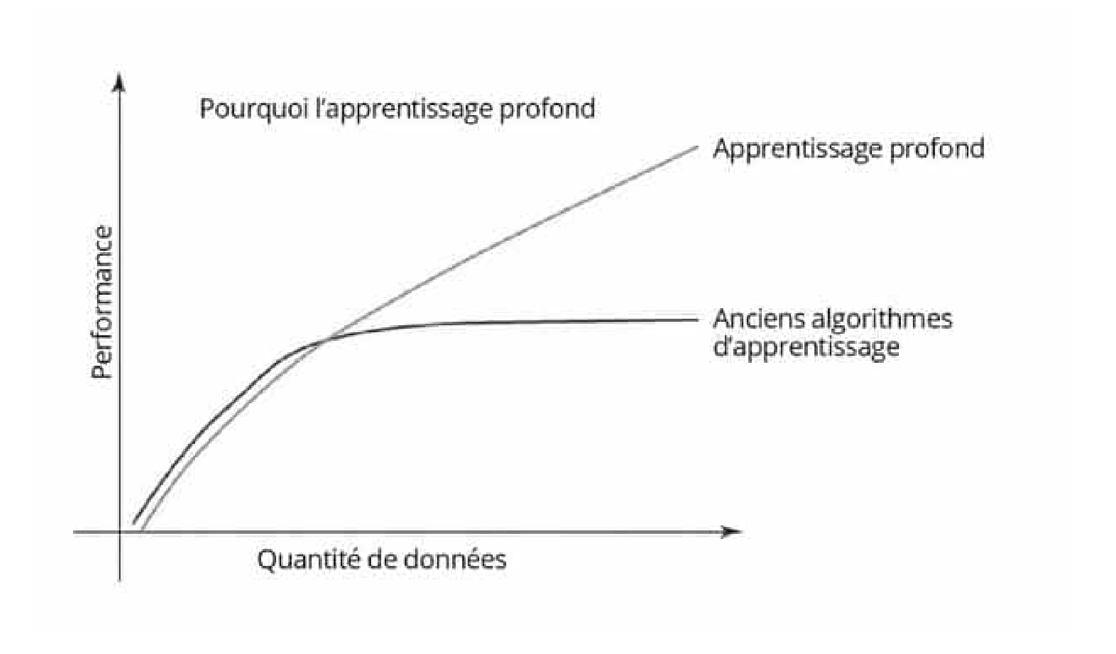


FIGURE 2.2 : Avec les solutions actuelles de l'IA, qui dit plus de données dit plus d'intelligence.

Ainsi, une photo de vos chatons devient de plus en plus utile, non pas simplement en raison de sa valeur affective – le fait qu'elle représente vos mignons petits compagnons à quatre pattes – mais parce qu'elle peut faire partie du processus d'apprentissage d'une IA, qui assimilera des concepts généraux comme les caractéristiques du chat ou la compréhension de ce que signifie être mignon.

À plus grande échelle, une compagnie comme Google alimente ses algorithmes avec des données en accès libre, comme le contenu des sites Internet ou les textes et les livres accessibles au public. Les « araignées »

logicielles de Google parcourent le Web, sautent d'un site à un autre et récupèrent des pages Internet avec les textes et les images qu'elles contiennent. Même si Google retourne aux utilisateurs une partie des données sous forme de résultats de recherche, elle extrait des données d'autres types d'informations à l'aide de ses algorithmes d'IA, lesquels apprennent ainsi à atteindre d'autres objectifs.

Des algorithmes qui traitent les mots peuvent permettre aux systèmes d'IA de Google de cerner et d'anticiper vos besoins, même si vous les exprimez non pas en utilisant une série de mots-clés, mais en parlant un langage normal pas nécessairement très clair, le langage parlé de tous les jours (et le langage de tous les jours n'est souvent pas clair). Si vous lancez une recherche sur Google en posant des questions au lieu d'écrire de simples chaînes de mots-clés, vous remarquerez que le moteur de recherche vous apporte souvent une réponse appropriée. Depuis 2012, avec l'apparition de la mise à jour Colibri (http://searchengineland.com/google-hummingbird-172816), Google est devenu mieux à même de comprendre les synonymes et les concepts, une chose qui va au-delà des données initiales acquises, et c'est là le résultat d'un processus d'IA. Il existe chez Google un algorithme plus sophistiqué encore, RankBrain, qui apprend directement à partir des millions de requêtes reçues chaque jour et qui est capable de répondre à des requêtes ambiguës ou peu claires, même si elles sont exprimées en termes argotiques ou familiers, ou truffées d'erreurs. RankBrain ne sert pas toutes les requêtes, mais il se sert des données pour apprendre à mieux leur répondre. Il prend déjà en charge 15 % des requêtes de Google, et il se pourrait bien qu'il en prenne en charge jusqu'à 100 % un jour ou l'autre.

Exploiter les données avec succès

Disposer d'une abondance de données ne suffit pas pour créer une IA performante. À l'heure actuelle, un algorithme d'IA ne peut pas extraire directement des informations des données brutes. La plupart des algorithmes ne fonctionnent que lorsque les données ont été recueillies et manipulées préalablement à leur analyse. Des informations utiles recueillies par un algorithme ne représentent pas nécessairement la bonne information. Les sections qui suivent vous aideront à comprendre comment recueillir et manipuler les données et comment automatiser leur collecte, grâce à une vue d'ensemble.

Quelles sont les sources des données ?

Les données que vous utilisez proviennent de diverses sources. Les sources de données les plus courantes sont les informations renseignées par les personnes humaines à un moment donné. Même lorsqu'un système collecte automatiquement des données sur un site d'achat, c'est d'abord l'utilisateur humain qui a saisi l'information : en cliquant sur tel ou tel article, en l'ajoutant à son panier, en spécifiant des caractéristiques (comme la taille ou la couleur) et une quantité, puis en fournissant ses coordonnées bancaires et postales. Par la suite, après l'achat, il évalue son expérience d'achat, le produit et le service de livraison, et écrit un commentaire. En résumé, toute expérience d'achat devient également un exercice de collecte de données.

Aujourd'hui, de nombreuses sources de données ont pour origine l'information recueillie auprès de personnes

humaines. Ce sont aussi les humains qui fournissent les intrants manuels. Pour rencontrer un professionnel, vous appelez un bureau ou vous vous y rendez. Un employé recueille auprès de vous les informations qui sont nécessaires pour le rendez-vous. Ces données collectées manuellement entrent ensuite dans une base de données quelque part, à des fins d'analyse.

Les données sont aussi collectées par des capteurs, qui peuvent prendre pratiquement n'importe quelle forme. Ainsi, par exemple, de nombreux organismes basent leur collecte de données physiques, comme le nombre de personnes qui visualisent un objet dans une fenêtre, sur la détection par le téléphone cellulaire. Les logiciels de reconnaissance faciale sont susceptibles de détecter les clients réguliers.

Cependant, les capteurs permettent de créer des séries de données à partir de pratiquement rien. La météo utilise des ensembles de données créés par des capteurs qui surveillent des conditions climatiques comme la pluie, la température, l'hygrométrie, la couverture nuageuse, *etc*. Des systèmes de contrôle robotisés permettent de corriger les petits ratés de l'activité robotique en analysant constamment les données recueillies par des capteurs. Associé à une petite application d'IA, un capteur peut vous annoncer le moment où votre dîner est prêt. Le capteur recueille des données, mais c'est l'application d'IA qui détermine, grâce à des règles, le moment où le plat est cuit à point.

Obtenir des données fiables

La notion de fiabilité semble facile à définir, mais elle est difficile à mettre en œuvre. Un processus est fiable quand les résultats qu'il produit sont ceux attendus et sont constants. Une source de données fiable produit des

données ordinaires qui ne réservent aucune surprise : personne n'est choqué le moins du monde par les données produites. Certes, selon le point de vue que vous adoptez, il peut être souhaitable que la plupart des gens ne bâillent pas et ne soient pas sur le point de s'endormir quand ils examinent les données. Ce sont souvent les surprises qui motivent l'analyse et l'examen des données. Par conséquent, une certaine dualité est associée aux données. D'un côté, nous voulons des données fiables, ordinaires et tout à fait prévisibles qui confirment ce que nous savons déjà, mais d'un autre côté, c'est en premier lieu dans l'imprévu que réside l'utilité de la collecte de données.

Néanmoins, il ne s'agit pas d'obtenir des données qui soient inhabituelles au point d'être alarmantes. Il convient de conserver un équilibre. Les données doivent demeurer dans certaines limites (comme l'explique la section « Adapter les données », plus loin dans ce chapitre). Elles doivent aussi répondre à des critères spécifiques pour être des valeurs de vérité (voir la section « Tenir compte des cinq types de données incorrectes », plus loin dans ce chapitre). Les données doivent aussi être obtenues en temps et en heure, et tous les champs des enregistrements de données entrantes doivent être remplis.



Dans une certaine mesure, la fiabilité des données est aussi conditionnée à leur sécurité. Leur cohérence revêt plusieurs aspects. Quand les données arrivent, on peut s'assurer qu'elles sont bien comprises dans les intervalles prévus et qu'elles se présentent sous une forme particulière. Cependant, une fois qu'elles sont stockées, leur fiabilité peut diminuer, sauf si l'on veille à ce qu'elles conservent leur forme attendue. Dès qu'une entité intervient sur ces données, leur fiabilité peut en souffrir et elles peuvent devenir douteuses ou potentiellement inutilisables dans une analyse ultérieure. Pour que les

données restent fiables, il faut qu'après leur arrivée personne ne les trafique en vue de les adapter à un certain domaine (par exemple en faisant en sorte qu'il n'y ait plus aucune donnée atypique).

Rendre la participation humaine plus fiable

Les humains font des erreurs, et l'erreur est humaine. Il ne serait pas raisonnable de s'attendre à ce qu'il n'y ait pas d'erreurs. Pourtant, un certain nombre d'applications sont conçues en supposant, d'une certaine manière, que les utilisateurs ne commettront aucune erreur d'aucune sorte. Il est prévu que chacun suivra simplement les règles. Malheureusement, dans la réalité, on peut être sûr qu'une vaste majorité des utilisateurs n'auront même pas lu les règles, car la plupart des gens sont trop paresseux, trop négligents ou trop pressés par le temps lorsqu'ils doivent faire des choses qui ne leur sont pas directement utiles.

Prenons l'exemple de la saisie d'un État américain dans un formulaire. Si l'on prévoit un simple champ de texte, certains utilisateurs saisiront le nom intégral : Kansas. Naturellement, certains utilisateurs feront une faute de frappe ou de typographie et saisiront par exemple Kansus, ou bien kANSAS. Les individus et les organisations n'exécutent pas tous la même tâche de la même manière. Dans le monde de l'édition et des médias, certains pourront se référer au guide de style de l'Associated Press (AP) et saisir Kan. Une personne de l'ancienne génération, habituée aux lignes directrices du gouvernement des États-Unis, saisira plutôt Kans. (avec un point). D'autres abréviations sont également utilisées. La poste américaine utilise le sigle KS, mais les garde-côtes américains utilisent le sigle KA. Quant à l'Organisation internationale de normalisation (ISO), elle

utilise l'abréviation US-KS. Or, il ne s'agit là que de la saisie du nom d'un État américain, une chose raisonnablement simple, du moins, c'est ce que l'on aurait pu penser. Sachant que ce nom ne devrait pas changer dans un avenir prévisible, on pourrait simplement faire apparaître dans le champ de saisie une liste déroulante afin que l'utilisateur choisisse la bonne entrée dans le format requis. Ainsi, ces différences d'abréviations, d'usage et de typographie et ces erreurs seraient éliminées d'un coup.



Les listes déroulantes sont adaptées à toutes sortes d'entrées de données, et elles rendent la saisie de données par un utilisateur humain extrêmement fiable, puisque l'utilisateur n'a pas d'autre possibilité que d'utiliser une des entrées proposées par le système. Bien sûr, il peut toujours choisir une entrée incorrecte, et c'est là qu'intervient le double contrôle. Certaines applications récentes comparent le code postal saisi au nom de ville saisi pour en vérifier la correspondance. En cas d'incohérence, l'utilisateur est invité à saisir un autre nom. La procédure peut paraître énervante, mais l'utilisateur n'y sera pas souvent confronté.

Même avec des contrôles et des entrées statiques, l'utilisateur a toujours la possibilité de faire des erreurs. La saisie de nombres, par exemple, peut être problématique. Lorsqu'il s'agit de saisir « 2,00 », ce peut être aussi « 2 », ou « 2,0 », ou encore « 2, » et il existe encore d'autres entrées possibles. Heureusement, l'entrée peut être analysée et reformatée, ce qui règle le problème, et cette tâche peut être exécutée automatiquement sans la participation de l'utilisateur.

Malheureusement, le reformatage ne peut pas corriger une entrée numérique erronée. On peut atténuer en partie l'erreur en prévoyant un contrôle de l'ordre de grandeur ou du signe. Un client ne pourra pas acheter – 5 pains de savon. S'il veut renoncer à son achat et renvoyer les

produits, il doit suivre une procédure de retour, et non une procédure d'achat. S'il s'agit simplement d'une erreur de saisie, on peut prévoir un message d'erreur indiquant l'intervalle dans lequel doit se trouver le nombre à saisir.

Recourir à la collecte de données automatisée

Certains pensent que la collecte de données automatisée résout tous les problèmes de saisie associés aux bases et jeux de données. En effet, elle offre un certain nombre d'avantages :

- davantage de cohérence ;
- davantage de fiabilité ;
- un moindre risque de données manquantes ;
- une plus grande exactitude ;
- une variance réduite pour les entrées programmées.

Malheureusement, il serait tout simplement incorrect de dire que la collecte de données automatisée résout tous les problèmes. Elle fonctionne au moyen de capteurs d'applications et de matériels informatiques qui sont conçus par des humains et qui ne fournissent que l'accès aux données que des humains décident de permettre. Compte tenu des limites que les humains imposent aux caractéristiques de la collecte de données automatisée, le résultat fournit souvent des informations moins utiles que ce qu'espéraient les concepteurs. Par conséquent, la collecte de données automatisée fait l'objet d'une évolution continue, à mesure que ses concepteurs s'emploient à résoudre les problèmes d'entrée de données.

La collecte de données automatisée souffre aussi des erreurs logicielles et matérielles qui peuvent être présentes dans tout système informatique, mais avec davantage de possibilités de *problèmes flous* (ce que l'on observe quand le système fonctionne bien en apparence mais ne produit pas le résultat désiré) que d'autres types de systèmes informatisés. Quand le système fonctionne, la fiabilité de l'entrée de données excède de loin les capacités humaines. Cependant, en cas de problème flou, le système, contrairement à l'humain, ne reconnaît souvent pas qu'il existe un problème, si bien que la base de données risque de contenir de plus en plus de données de qualité médiocre, ou même, de mauvaises données.

Adapter les données

L'emploi du mot *manipulation* à propos des données peut suggérer que les données sont modifiées de façon sournoise ou peu scrupuleuse. Peut-être devrait-on plutôt dire que les données sont adaptées, ou *mises en forme*, ce qui suggère qu'on les rend plus belles, plus présentables. Quel que soit le terme utilisé, des données brutes répondent rarement aux exigences du traitement et de l'analyse. Pour pouvoir tirer quelque chose des données, il faut les adapter à des besoins spécifiques. Les sections qui suivent traitent de ces besoins et de ces exigences relatives aux données.

Faire avec des données manquantes

Pour pouvoir répondre correctement à une question, il faut connaître tous les faits. Autrement, on peut se livrer à des suppositions et « sauter aux conclusions » sans se fonder sur tous les éléments nécessaires, mais dans ce cas, la

réponse que l'on aura choisie aura autant de chances d'être incorrecte que d'être la bonne réponse. En analysant des données, il vous est probablement arrivé, plus souvent que vous pourriez le croire, de sauter aux conclusions alors que vous ne disposiez pas de toutes les données. Un *enregistrement* ou *bloc de données*, c'est-à-dire une entrée de la *base de données* (l'ensemble des données), est constitué de champs contenant des éléments qui permettent de répondre à une question. Chaque champ contient un type unique de donnée qui représente un fait. Quand ce champ est vide, vous ne disposez pas des données nécessaires pour pouvoir répondre à la question sur la base de cet enregistrement de données.



Dans le cadre du processus de traitement des données manquantes, il importe que vous sachiez que ces données sont manquantes. Or, il est parfois très difficile de déterminer qu'une information est manquante dans un ensemble de données, sachant qu'il faut, pour cela, examiner les données à un niveau élémentaire, une chose que la plupart des gens ne sont pas préparés à faire et qui demande du temps, même quand on a les compétences requises. Souvent, le premier indice d'une donnée manquante est la réponse absurde que l'algorithme et la base de données apportent à votre question. Si l'algorithme est correct, alors le problème est dans les données.

Un problème peut se poser quand le processus de collecte de données n'inclut pas toutes les données nécessaires pour répondre à une question particulière. Parfois, il vaut mieux faire carrément abstraction d'une information plutôt que de prendre en compte une donnée considérablement altérée. Si vous constatez que dans un champ particulier, au moins 90 % des données sont manquantes, ce champ devient inutile et il convient de

l'éliminer de la base de données (ou de trouver un moyen d'obtenir toutes ces données).

Dans les situations moins extrêmes, on peut distinguer deux cas. Lorsque l'omission de données présente un caractère aléatoire, c'est souvent le résultat d'une erreur de l'homme ou du capteur. C'est ce qui se produit quand certaines entrées sont manquantes çà et là dans une série de données. Un simple dysfonctionnement ou un simple bogue peut en être la cause. Quand ce sont des données successives qui manquent, il s'agit d'une défaillance généralisée. Si les informations requises sont absentes sur tout un segment des enregistrements, c'est toute l'analyse qui risque d'être biaisée.

Lorsque l'absence d'une donnée est aléatoire, il est très facile d'y remédier. On peut remplacer la donnée manquante par une simple valeur médiane ou moyenne. Les données ne seront pas entièrement exactes, mais l'approximation permettra d'obtenir une réponse raisonnablement valable. Dans certains cas, les spécialistes utilisent un algorithme spécial pour calculer la valeur manquante, ce qui peut rendre l'ensemble de données plus adéquat aux dépens du temps de traitement.

En cas d'absence de plusieurs données successives, le problème est bien plus difficile, voire impossible à résoudre, puisqu'il n'est pas possible de supposer une valeur d'après les données au milieu desquelles elle se trouve. Quand on trouve la cause de cette absence de données, on peut parfois reconstituer les données manquantes. Quand la reconstitution est impossible, on peut décider d'ignorer le champ concerné. Malheureusement, ce champ est nécessaire pour certaines réponses, et il faudra alors ignorer cette série particulière d'enregistrements, ce qui risque d'amener un résultat incorrect.

Prendre en compte une discordance entre les données

Il se peut que tous les enregistrements d'un jeu de données soient dûment renseignés, mais qu'il y ait une mauvaise harmonisation entre ces données et celles provenant d'un autre jeu de données que vous utilisez. Ainsi, par exemple, les données d'un champ numérique peuvent être des nombres à virgule flottante dans un jeu de données, et des nombres entiers dans un autre. Or, avant de pouvoir combiner les deux jeux de données, il faudra que les deux champs contiennent le même type de données.

Il peut exister des discordances de toutes sortes. Par exemple, on sait que les champs de date peuvent être sous des formats variés. Or, pour pouvoir comparer deux dates, il faut que le format de date soit le même. Un autre problème avec les dates est qu'elles peuvent paraître identiques sans l'être, par exemple si les dates d'un jeu de données se réfèrent à l'heure moyenne de Greenwich (GMT) tandis que celles d'un autre jeu de données font référence à un autre fuseau horaire. Avant de pouvoir comparer les dates ou les heures, il faut alors convertir les dates au même fuseau. La situation peut encore se compliquer lorsque les dates d'un seul des deux jeux de données proviennent d'un site qui utilise l'heure d'été.

Même quand le type et le format de date diffèrent, d'autres discordances peuvent se produire. Il se peut, par exemple, que les champs d'un jeu de données ne soient pas en phase avec ceux de l'autre. Dans certains cas, ces différences sont faciles à corriger. Il arrive que le prénom et le nom de famille, dans un jeu de données, constituent une donnée unique dans un seul champ, tandis que dans un autre jeu de données, on aura utilisé deux champs distincts. La solution consiste alors à modifier les jeux de

données de manière à ce que tous utilisent un champ unique, ou au contraire à remplacer le champ unique, partout où il existe, par deux champs distincts, un pour le nom et un autre pour le prénom. Malheureusement, un certain nombre de discordances dans les données posent davantage de problèmes, et il est tout à fait possible que vous ne puissiez pas du tout les identifier. Toutefois, avant de baisser les bras, réfléchissez à ces solutions éventuelles :

» Calculer les données manquantes en se référant à d’autres données disponibles.

» Localiser les données manquantes dans un autre jeu de données.

» Combiner les jeux de données pour former un jeu global avec des champs cohérents.

» Recueillir des données supplémentaires à partir d’autres sources pour remplacer les données manquantes.

» Redéfinir la question de telle sorte que les données manquantes ne soient plus nécessaires.

Séparer les données utiles des autres données

Dans certaines organisations, on considère qu’il n’y a jamais trop de données. Or, l’excès de données peut devenir un problème tout aussi important que le manque de données. Pour qu’une IA puisse résoudre efficacement les problèmes, il faut qu’elle dispose de la bonne quantité de données, ni plus ni moins. Il est utile de définir de façon claire et concise la question à laquelle il s’agit de répondre, et d’utiliser le bon algorithme (ou la bonne série d’algorithmes). Naturellement, les principaux problèmes

que pose l'excès de données sont l'allongement du temps nécessaire pour trouver la solution (après être venu à bout de toutes ces données pléthoriques), et le risque d'obtenir des résultats confus ou déroutants, lorsque l'arbre cache la forêt.



Pour créer le jeu de données que vous devez analyser, ne touchez pas aux données originales mais faites-en une copie. Conservez toujours les données brutes originales, sans rien y changer, afin de pouvoir les exploiter par la suite dans le cadre d'une autre étude. Par ailleurs, obtenir un jeu de données qui convienne pour l'analyse nécessite parfois plusieurs essais successifs, car il se peut que vous constatiez que le résultat de votre requête ne répond pas à vos besoins. Il s'agit de créer un jeu de données qui ne contienne que les données nécessaires à l'analyse, mais en gardant à l'esprit qu'il vous faudra peut-être soumettre ces données à une forme spécifique de formatage pour pouvoir obtenir le résultat désiré.

Tenir compte des cinq types de données incorrectes

Nous avons l'habitude de considérer l'information comme ce qu'elle est effectivement dans bien des cas : une opinion. En vérité, il arrive que les données soient faussées au point qu'elles deviennent inexploitables. L'ordinateur ne peut pas faire la différence entre une donnée fiable et une donnée biaisée : pour lui, une donnée est une donnée, c'est tout. Un des problèmes qui font qu'il est difficile, sinon impossible, de créer une IA qui raisonne réellement comme un être humain est que si les humains peuvent gérer les contre-vérités, les ordinateurs en sont incapables. Le mieux que l'on puisse espérer est que les données faussées soient traitées comme des anomalies et ne soient donc pas prises en compte, mais

cette technique ne résout pas nécessairement le problème, car une personne utilisera tout de même les données et essaiera de déterminer la vérité à partir des informations fausses qui s'y trouvent.



Dans le souci de disposer de jeux de données moins corrompus, on pense souvent qu'au lieu de compter sur des personnes pour saisir les données, il doit être possible de faire recueillir les données par des capteurs et autres systèmes. Seulement, les capteurs et autres méthodologies d'obtention de données reflètent les objectifs de leurs inventeurs humains et les limites à ce qu'une technologie particulière est capable de détecter. Par conséquent, même des données provenant d'une machine ou d'un capteur ne seront pas nécessairement exemptes de faussetés qu'une IA pourra difficilement détecter.

Dans les sections suivantes, le principal exemple servant à illustrer les cinq types de fausseté pouvant apparaître dans les données est celui d'un accident de la route. Les concepts que l'on se propose d'aborder à travers cet exemple ne seront pas toujours présents dans les données, et ils pourront apparaître selon des modalités différentes de celles étudiées. Il n'en demeure pas moins qu'il est normalement nécessaire de tenir compte de ce genre de problème quand on examine des données.

Les données mensongères

Les données mensongères sont celles qui reflètent une tentative délibérée de remplacer une information vraie par une information fausse. Dans un constat d'accident, par exemple, un automobiliste pourra affirmer avoir été momentanément aveuglé par le soleil, si bien qu'il lui aurait été impossible de voir l'obstacle. En réalité, il aura été distrait par quelque chose, ou par ses pensées. Si

personne n'est en mesure de réfuter cette affirmation, il aura réussi à mieux tirer son épingle du jeu. Surtout, la compagnie d'assurance calculera les primes en se fondant sur des données faussées.



On pourrait penser qu'il est toujours possible d'éviter la présence de données mensongères, mais ce n'est souvent pas le cas. Les gens disent des « pieux mensonges » pour laisser quelqu'un sauver la face ou pour résoudre un problème en faisant le moins d'efforts possible. Une donnée fausse peut aussi provenir d'une rumeur. En réalité, les sources des erreurs de ce type sont si nombreuses qu'il est vraiment difficile d'imaginer qu'il soit possible de toujours les éviter. Cela dit, il s'agit tout de même d'un type d'information fausse qu'il est possible d'éviter le plus souvent.

Les omissions volontaires

Il y a omission volontaire quand une personne dit la vérité tout en évitant de révéler un fait important qui changerait la perception de la situation dans son ensemble. Pour reprendre l'exemple d'un accident de la route, supposons qu'un automobiliste ait percuté un cerf et que sa voiture ait subi d'importants dégâts. Il déclare à juste titre que la route était mouillée et qu'il faisait sombre, qu'il n'a peut-être pas freiné assez vite, et que l'animal a surgi d'un fourré situé au bord de la route. Jusqu'ici, on pourrait conclure que c'est simplement un accident, par malchance. En réalité, l'automobiliste se garde de dire qu'au moment où l'animal a surgi, il était en train d'écrire un SMS. Si les gendarmes disposaient de cette information, ils retiendraient comme cause de l'accident l'inattention du conducteur. Celui-ci serait passible d'une amende et l'expert des assurances ne renseignerait pas la base de données de la même manière. Ici encore, le

calcul des primes par l'assureur sera fondé sur des données faussées.



Il est pratiquement impossible d'éviter les données mensongères et les omissions volontaires. Quelqu'un peut exclure délibérément certains faits dans un rapport, mais quelqu'un peut tout aussi bien oublier d'y inclure certains faits. Après un accident, on est généralement secoué et on a vite fait de perdre de vue certains éléments et de ne mentionner que ceux qui ont laissé l'impression la plus forte. Même si la personne en question se remémore par la suite certains détails et en fait état, il sera peu probable que la base de données contienne tous les éléments de la vérité.

Les erreurs de perspective

Les erreurs de perspective se produisent quand plusieurs parties ont un point de vue différent sur la situation. Considérons le cas d'une collision entre une automobile et un piéton. L'automobiliste, le piéton renversé et le témoin de l'accident, même en supposant que chacun sache et dise la vérité, exprimeront vraisemblablement des points de vue divergents. L'expérience montre que c'est pratiquement toujours ainsi que les choses se passent, et l'agent de police qui aura recueilli la version de chacun devra trouver un moyen terme entre les différentes déclarations, tout en y mêlant sa propre expérience. En d'autres termes, son rapport sera proche de la vérité, mais pas suffisamment proche pour les besoins d'une IA.

Il est important de tenir compte de ces divergences de points de vue. L'automobiliste pouvait voir les indications de son tableau de bord au moment de l'accident, or les deux autres parties ne disposent pas de ces informations. Quant au piéton renversé, il a pu voir l'expression du visage du conducteur (et connaître ainsi son intention). Le

témoin, lui, était sans doute le mieux placé pour savoir si l'automobiliste a bien tenté de s'arrêter, ou s'il a dévié ou non de sa trajectoire. Chacune des parties relatera les faits selon ce qu'elle a vu, mais certaines données ne lui seront pas connues.



La divergence entre les points de vue est peut-être la forme la plus redoutable d'inexactitude, car quiconque essaiera de déterminer la vérité ne pourra finalement obtenir qu'une version intermédiaire, dans le meilleur des cas, et ce ne sera jamais une version exacte des faits. Un être humain qui visualise l'information peut compter sur son intuition ou sur son instinct pour parvenir à une meilleure approximation de la vérité, mais une IA n'utilisera jamais que la version intermédiaire, ce qui signifie que l'IA est toujours significativement désavantagée. Malheureusement, il est impossible d'éviter les biais liés aux divergences de points de vue car quel que soit le nombre de témoignages dont on dispose, on ne peut espérer obtenir qu'une approximation de la vérité, et non pas la pure vérité.

Il existe aussi une autre sorte d'inexactitude à prendre en compte, liée à une question de perspective. Remontons à l'année 1927 et prenons le cas d'un sourd qui se rendrait chaque semaine au cinéma pour voir un film muet. Durant une heure, le temps de la projection, il aurait l'impression d'être comme tout le monde. En effet, son expérience serait exactement la même que celle de n'importe qui d'autre. En octobre de la même année, le cinéma annoncerait l'installation d'un système sonorisé en vue de la projection de films parlants. D'après l'annonce, ce serait une formidable nouveauté. Tout le monde en conviendrait, sauf cet homme sourd qui aurait soudain l'impression d'être un citoyen de seconde classe, différent des autres, et d'une certaine manière, exclu. De son point de vue, l'annonce serait mensongère : la sonorisation serait la pire chose qui soit, et non la meilleure. Ce qui paraît

généralement vrai ne l'est pas pour tout le monde. L'idée d'une vérité universelle est un mythe. Une vérité universelle, cela n'existe pas.

Les biais

On parlera de biais quand une personne est en mesure de connaître la vérité, mais incapable de vraiment la percevoir en raison de préoccupations ou de croyances qui lui sont propres. Pour reprendre l'exemple d'accident qui précède, supposons que l'automobiliste concentre son attention sur le milieu de la route au point que le cerf surgissant au bord du fourré soit pour lui complètement invisible. Il n'aura donc pas le temps de réagir quand l'animal décidera soudain de bondir au milieu de la route pour la traverser.

Un problème qui se pose avec le biais est qu'il peut être particulièrement difficile à catégoriser. Si l'automobiliste n'a pas vu le cerf, on peut parler d'un véritable accident dans la mesure où l'animal était caché par la végétation. Cependant, il se peut aussi que le conducteur soit coupable d'inattention, dans la mesure où il ne dirigeait pas son attention sur ce qu'il fallait. Il peut aussi avoir été momentanément distrait. En un mot, la question n'est pas de savoir s'il a vu ou non l'animal, mais pourquoi il ne l'a pas vu. Dans bien des cas, quand on conçoit un algorithme destiné à éviter une source de biais, il devient important de confirmer cette source.



En théorie, il est toujours possible d'éviter les biais. Cependant, dans la réalité, tous les êtres humains manifestent des biais de divers types, et ces biais entraînent inévitablement la présence de données faussées. Obtenir d'une personne qu'elle fixe son attention sur un objet et qu'elle l'enregistre tel quel dans sa tête n'est pas une mince affaire. Afin d'éviter une surcharge

d'informations, l'être humain utilise des filtres, et ces filtres sont aussi une source de biais car ils l'empêchent de voir vraiment les choses.

Le cadre de référence

Parmi les cinq types de données fausses, le problème du cadre de référence n'est pas le résultat d'une erreur quelconque, mais d'un problème de compréhension. On rencontre ce problème quand une des parties en présence décrit une situation ou un événement, par exemple un accident, et que l'autre partie, n'ayant pas l'expérience de cette situation ou de cet événement, en a une perception confuse, ou même totalement fausse. Les erreurs liées au cadre de référence sont un ingrédient abondamment utilisé dans les comédies. Un exemple célèbre est le sketch d'Abbott et Costello *Who's On First ?* (https://www.youtube.com/watch?v=kTcRRaXV-fg). Il peut être impossible à une personne de comprendre ce que lui dit son interlocuteur lorsqu'il lui manque la connaissance de la situation ou du contexte, c'est-à-dire le cadre de référence.

Un autre exemple est celui de deux parties qui ne peuvent absolument pas se comprendre. Supposons qu'un marin ait affronté une dangereuse tempête en mer. Même avec des vidéos, des interviews ou un simulateur, il sera impossible de faire partager cette expérience à une personne n'ayant jamais vécu l'expérience d'une tempête, même sur la terre ferme, car cette personne n'a aucun cadre de référence.



Le meilleur moyen d'éviter les problèmes de cadre de référence consiste à faire en sorte que toutes parties concernées développent des cadres de référence similaires. Pour cela, il faut que les différentes parties acquièrent les mêmes connaissances expérientielles.

Ainsi, le transfert de données entre deux personnes pourra être pertinent. Cependant, quand on travaille avec un jeu de données, qui sont nécessairement enregistrées et statiques, des erreurs liées au cadre de référence peuvent toujours se produire, si l'utilisateur potentiel n'a pas la connaissance expérientielle nécessaire.

Dans le cas d'une IA, des problèmes de cadre de référence se poseront toujours car une IA ne peut pas créer une expérience. Il en va tout autrement d'une banque de données de connaissances acquises. La banque de données contient des faits, or l'expérience se base non seulement sur des faits, mais aussi sur des conclusions que les technologies actuelles ne sont pas capables de reproduire.

Définir les limites de l'acquisition de données

Vous avez peut-être l'impression que les gens acquièrent vos données sans rime ni raison, et vous voyez juste. Effectivement, des organisations recueillent, catégorisent et stockent des données sur chacun de nous, apparemment sans but particulier. D'après Data Never Sleeps (https://www.domo.com/blog/data-never-sleeps-5/), des données sont collectées partout dans le monde au rythme de 2,5 quintillions d'octets par jour. Ces données quotidiennes peuvent prendre des formes très variées, comme en attestent les exemples suivants :

- » Google effectue 3607080 recherches.
- » Les utilisateurs de Twitter envoient 456000 tweets.
- » Les utilisateurs de YouTube regardent 4146600 vidéos.

» Les boîtes de réception reçoivent 103447529 courriers électroniques indésirables.
» The Weather Channel reçoit 18055555,56 requêtes concernant la météo.
» GIPHY distribue 694444 images au format GIF.

Dans les entreprises et autres organisations du monde entier, l'acquisition de données est devenue une drogue, et c'est à croire que celle qui en acquerra le plus gagnera un prix. Pourtant, en elle-même, l'acquisition de données ne produit rien du tout. Le *Guide du voyageur galactique*, de Douglas Adams (https://www.amazon.fr/H2G2-Guide-voyageur-galactique/dp/2070437434), illustre clairement ce problème. Dans ce livre, des supercréatures fabriquent un ordinateur immense pour calculer le sens de « la vie, l'univers, et tout ». La réponse proposée, le nombre 42, ne résout rien, aussi certaines créatures se plaignent que la collecte, la catégorisation et l'analyse de toutes les données utilisées pour formuler la réponse, n'aient pas permis d'obtenir un résultat exploitable. L'ordinateur, qui est doué de conscience de soi, dit aux utilisateurs que la réponse est correcte, or il leur manque de connaître la question pour que cette réponse ait un sens. L'acquisition de données peut se faire en quantités illimitées, mais trouver les bonnes questions à poser est parfois très rebutant, voire impossible.



Le principal problème qu'une organisation doit résoudre en matière d'acquisition de données est de savoir quelles questions poser et pourquoi ces questions sont importantes. Il importe d'adapter l'acquisition de données aux questions auxquelles il faudra répondre. Si vous gérez un commerce en ville, par exemple, les questions auxquelles vous devrez apporter une réponse pourront être les suivantes :

» Combien de personnes passent chaque jour devant votre boutique ?

» Combien s'arrêtent devant votre vitrine ?

» Pendant combien de temps les gens regardent-ils ce que vous avez en vitrine ?

» À quel moment de la journée regardent-ils ce que vous avez en vitrine ?

» Est-ce que certaines présentations produisent de meilleurs résultats ?

» Lesquelles de ces présentations incitent les gens à entrer et à faire des achats ?

Cette liste pourrait s'allonger, mais l'idée est qu'il est essentiel de créer une liste de questions qui concernent vos besoins commerciaux. Ensuite, il faut vérifier que chacune de ces questions est réellement importante, qu'elle répond à un besoin, puis déterminer précisément les informations qu'il vous faut pour pouvoir répondre aux questions.



Naturellement, il serait impossible de recueillir toutes ces données à la main, et c'est là que l'automatisation entre en jeu. L'automatisation devrait permettre d'avoir des entrées de données fiables, reproductibles et cohérentes. Cependant, de nombreux facteurs intervenant dans l'automatisation de l'acquisition de données peuvent produire des données qui ne seront pas particulièrement utiles. Considérons, par exemple, les problèmes suivants :

» Les capteurs ne peuvent recueillir que les données qu'ils sont conçus pour recueillir. Si vous en faites une utilisation pour laquelle ils ne sont pas prévus, vous risquez de vous retrouver avec des données manquantes.

- » Les gens produisent plusieurs types de données fausses (voir la section « Tenir compte des cinq types de données incorrectes », précédemment dans ce chapitre), ce qui signifie que les données que vous recevez peuvent être fausses.
- » Les données peuvent être altérées quand les conditions de leur collecte ne sont pas définies correctement.
- » Si des données sont interprétées de façon incorrecte, les résultats seront également incorrects.
- » La transcription d'une question sous forme d'un algorithme compréhensible pour l'ordinateur est un processus au cours duquel des erreurs peuvent être générées.

Bien d'autres problèmes doivent être pris en compte (de quoi remplir un livre). Si vous combinez des données mal recueillies et mal formatées avec des algorithmes qui ne répondent pas vraiment à vos questions, le résultat que vous obtiendrez pourra conduire votre entreprise dans la mauvaise direction. C'est la raison pour laquelle on reproche souvent à l'IA de produire des résultats incohérents ou peu fiables. Pour que l'acquisition de données soit le genre d'outil sur lequel vous pouvez valablement vous appuyer, il faut poser la bonne question, obtenir les données correctes, exécuter le traitement adéquat, puis analyser correctement les données.

Chapitre 3

Réfléchir à l'utilisation d'algorithmes

DANS CE CHAPITRE…

» **Découvrir le rôle des algorithmes dans l'IA**

» **Gagner des parties avec la recherche d'espaces d'états et la méthode du minimax**

» **Analyser le fonctionnement des systèmes experts**

» **Comprendre que l'apprentissage machine et l'apprentissage profond font partie de l'IA**

Les données jouent un rôle capital pour l'IA. Les récentes avancées de l'IA indiquent que pour certains problèmes, choisir la bonne quantité de données est plus important que choisir le bon algorithme. Ainsi, en 2001, deux chercheurs de chez Microsoft, Banko et Brill, dans leur article mémorable intitulé « Scaling to Very Very Large Corpora for Natural Language Disambiguation » (http://www.aclweb.org/anthology/P01-1005), ont montré que pour créer un modèle de langage à l'aide d'un ordinateur, il n'était pas nécessaire de disposer de l'algorithme le plus intelligent qui soit. Une fois que l'on a soumis plus d'un milliard de mots en contexte pour la résolution d'un problème, n'importe quel algorithme commencera à être très performant. Ce chapitre vous aide à comprendre la relation entre les algorithmes et les

données utilisées pour leur faire accomplir des tâches utiles.

Cependant, quelle que soit la quantité de données dont vous disposez, il vous faut un algorithme afin de pouvoir les exploiter. Pour que ces données produisent quelque chose de correct lorsque vous utilisez les algorithmes choisis, vous devez aussi procéder à une *analyse des données* (une série d'étapes bien définissables). Vous ne pouvez pas y couper. Même si l'IA est une automatisation intelligente, l'automatisation doit parfois passer après l'analyse. Des machines qui apprennent d'elles-mêmes, qui savent discerner ce qui est approprié et ce qui ne l'est pas et qui rendent inutile toute intervention humaine, cela n'existera que dans un avenir lointain. La seconde partie de ce chapitre vous aide à comprendre le rôle des systèmes experts, de l'apprentissage machine, de l'apprentissage profond et d'applications comme AlphaGo dans les avancées qui permettent de rapprocher un petit peu plus de la réalité certaines possibilités futures.

Comprendre le rôle des *algorithmes*

Les gens reconnaissent généralement l'IA quand un outil représente une nouvelle approche et interagit avec l'utilisateur à la façon dont le ferait un être humain. C'est le cas avec des assistants numériques comme Siri, Alexa et Cortana. Cependant, certains autres outils d'utilisation courante comme les routeurs GPS et les systèmes spécialisés de pilotage automatique (comme ceux utilisés pour éviter les collisions entre les automobiles, pour piloter automatiquement les avions ou pour gérer les plans de production) ne font même pas penser à l'IA parce qu'ils sont très communément utilisés et considérés comme des choses banales, sachant qu'ils agissent en coulisses.

Il s'agit là clairement de l'*effet IA*, appelé ainsi et décrit par Pamela McCorduck, une auteure américaine qui a publié en 1979 un remarquable historique de l'intelligence artificielle. Le principe de l'effet IA est que les logiciels intelligents et performants sont rapidement négligés par le public et deviennent des acteurs silencieux, tandis que l'attention se porte sur les problèmes non encore résolus que pose l'IA. Les gens sont inconscients de l'importance des algorithmes classiques pour l'IA et se mettent à fantasmer sur une IA issue d'une technologie ésotérique, ou assimilent l'IA à de récentes avancées comme l'apprentissage machine et l'apprentissage profond.

Un *algorithme* est une procédure, c'est-à-dire une succession d'opérations, généralement exécutée par un ordinateur, qui garantit l'aboutissement à la solution correcte d'un problème dans un temps fini, ou qui vous dit qu'il n'existe aucune solution. Cela fait plusieurs millénaires que l'on résout manuellement des algorithmes, ce qui peut prendre énormément de temps et nécessiter de nombreux calculs, selon la complexité du problème à résoudre. Tout algorithme sert à trouver une solution, et les meilleurs algorithmes sont les plus simples et les plus rapides. Les algorithmes sont codés en dur dans l'intelligence des humains qui les ont conçus, et toute machine qui fonctionne en exécutant ces algorithmes ne peut que refléter l'intelligence incorporée dans ces procédures algorithmiques.

Comprendre le sens du mot algorithme

Un algorithme est toujours constitué d'une série d'étapes, mais il ne les exécute pas nécessairement toutes pour résoudre un problème. Le champ des algorithmes est extrêmement vaste. Les opérations qui les constituent

peuvent consister à stocker des données, à les explorer, à les classer ou à les ordonner sous forme de structures de données. Il existe des algorithmes pour résoudre des problèmes dans les domaines de la science, de la médecine, de la finance, de la production industrielle et des approvisionnements, et de la communication.

Tout algorithme est une série d'opérations pour trouver la solution correcte d'un problème dans un délai raisonnable (ou pour aboutir à la conclusion qu'aucune solution n'est trouvée). La différence entre les algorithmes de l'IA et les algorithmes génériques est qu'avec l'IA, la résolution des problèmes est considérée comme étant généralement (ou même exclusivement) le produit d'un comportement humain intelligent. Les algorithmes de l'IA servent généralement à résoudre des problèmes compliqués, qui font souvent partie des problèmes de classe NP, ou problèmes NP-complets (NP signifie non déterministe en temps polynomial) auxquels les humains s'attaquent en associant leur intuition à une approche rationnelle. En voici quelques exemples :

- les problèmes de planification et l'allocation des ressources rares ;
- la recherche de chemins dans des espaces physiques ou figuratifs complexes ;
- la reconnaissance des tendances dans la vision d'images (par opposition avec la restauration ou le traitement d'images) ou la perception des sons ;
- le traitement du langage (compréhension et traduction de textes) ;
- les jeux non compétitifs.



Les problèmes NP-complets se distinguent des autres problèmes algorithmiques par le fait qu'il ne soit pas encore possible de leur trouver une solution dans un délai

raisonnable. Ce ne sont pas des problèmes que l'on résout en essayant toutes les combinaisons possibles ou toutes les possibilités. Même avec des ordinateurs plus puissants que ceux dont nous pouvons disposer aujourd'hui, la recherche de la solution durerait éternellement, ou presque. De façon similaire, dans l'IA, ces problèmes sont qualifiés de *IA-complets*.

Commencer par un plan et des ramifications

La planification vous permet de déterminer la séquence d'actions à exécuter pour atteindre un certain objectif. C'est un problème d'IA classique, et l'on en trouve des exemples dans la production industrielle, l'allocation des ressources et le déplacement d'un robot dans une pièce. À partir de l'état présent, une IA détermine les actions possibles. Techniquement, elle change l'état courant en un certain nombre d'états futurs. Ensuite, elle change tous les états futurs en leurs états futurs, et ainsi de suite. Quand les états ne peuvent plus être changés et quand l'IA cesse ce développement, elle a créé un *espace d'états*, qui est constitué de tout ce qui peut se produire dans l'avenir. Une IA peut exploiter un espace d'états non pas simplement en produisant une éventuelle prédiction (en réalité elle prédit tout, quoique certains états futurs soient plus probables que d'autres), mais aussi en explorant les décisions qu'elle peut prendre pour atteindre son objectif de la meilleure manière possible. C'est ce que l'on appelle la *recherche d'espaces d'états*.

Travailler sur un espace d'états suppose l'utilisation de structures de données et d'algorithmes particuliers. Les structures de données principalement utilisées sont les arborescences et les graphes. Les algorithmes de prédilection pour parcourir les graphes de façon

performante sont l'algorithme de parcours en largeur (BFS) et l'algorithme de parcours en profondeur (DFS).

La construction d'une arborescence est similaire au développement d'un arbre dans le monde physique. Chaque élément que l'on ajoute est un *nœud*. Les nœuds sont reliés par des liens. Les nœuds et les liens forment ensemble une structure en forme d'arbre (voir Figure 3.2).



Les arborescences, ou arbres, ont un nœud racine, à l'image des arbres du monde réel. Le *nœud racine* est le point de départ pour le traitement à exécuter. À ce nœud racine sont reliées des branches ou des feuilles. Un nœud feuille est une terminaison de l'arbre. Les nœuds branches supportent d'autres branches ou bien des feuilles. Le type d'arbre de la Figure 3.1 est un arbre binaire, car chaque nœud a au maximum deux connexions (cependant, les arbres qui représentent des espaces d'états peuvent avoir plusieurs branches à chaque nœud).

Sur l'arborescence de la Figure 3.1, la branche B est le *fils* du nœud racine. En effet, le nœud racine apparaît en premier dans la liste. Les feuilles E et F sont deux fils de la branche B, si bien que la branche B est le *père* des feuilles E et F. La relation entre les nœuds est importante car dans l'étude des arborescences, on s'intéresse souvent aux relations fils/père entre les nœuds. Si l'on n'utilisait pas ces termes, les discussions relatives aux arborescences risqueraient d'être assez déconcertantes.

Un graphe est une sorte d'extension d'un arbre. Le graphe comporte aussi des nœuds qui sont reliés les uns aux autres pour constituer des relations. Cependant, contrairement aux arbres binaires, un nœud dans un graphe peut avoir plus d'une ou deux connexions. Les nœuds peuvent avoir une multitude de connexions, et, plus important, les relations peuvent exister dans les deux sens, pas seulement d'un père vers un fils. Pour éviter de

compliquer les choses, examinons le graphe simple de la Figure 3.2.

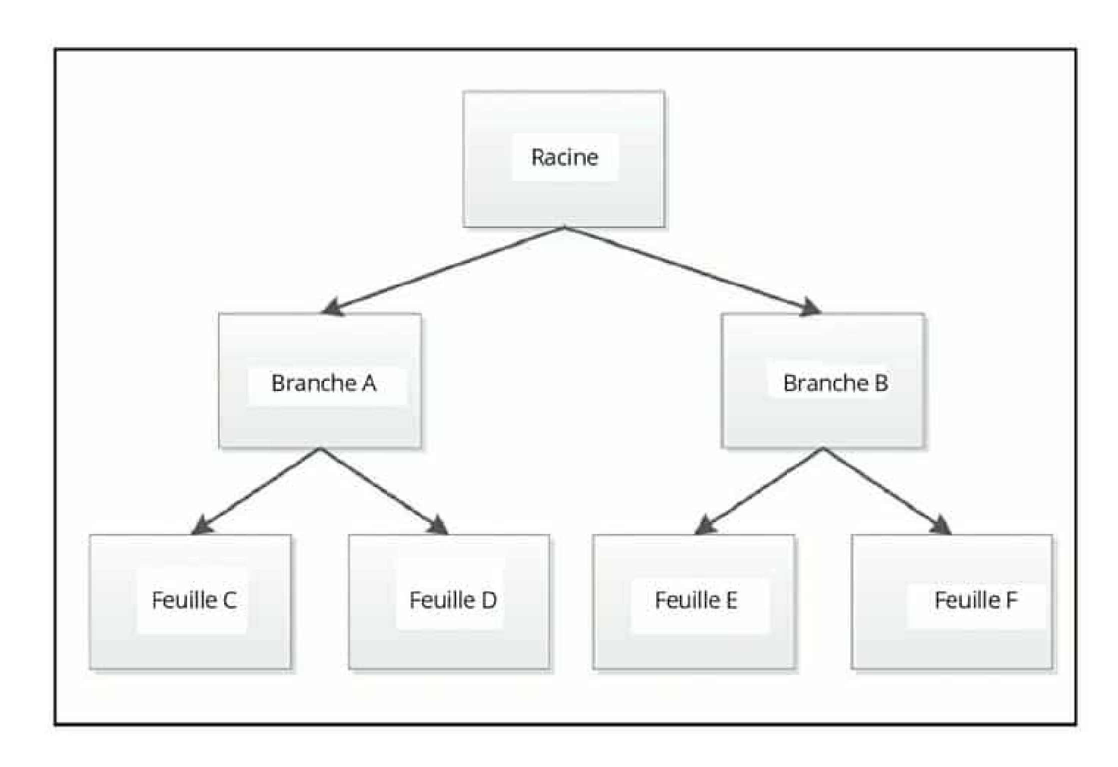


FIGURE 3.1 : Ici, un arbre peut aussi avoir ses racines dirigées vers le haut.

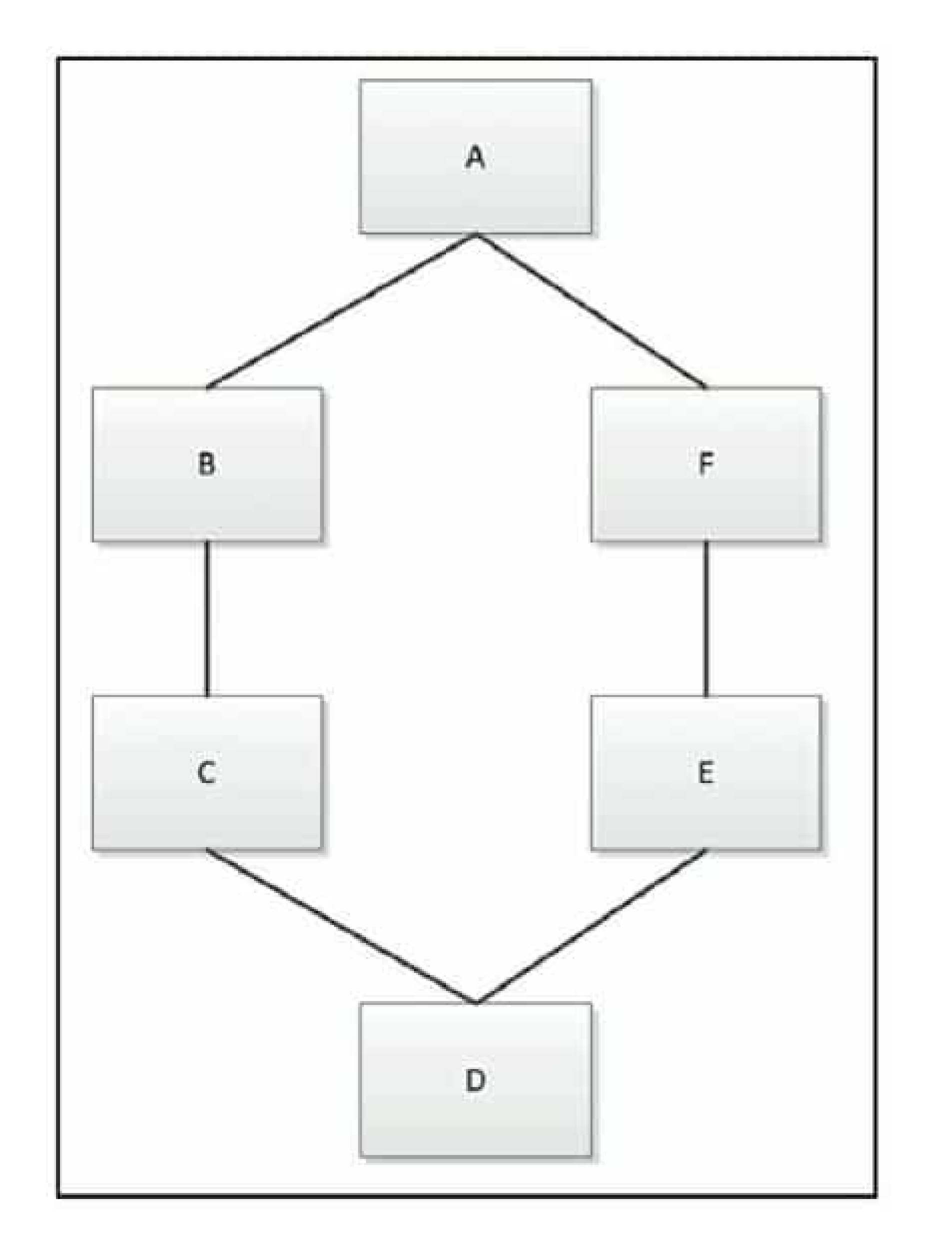


FIGURE 3.2 : Les nœuds d'un graphe peuvent être reliés de toutes sortes de manières.

Les graphes sont des structures qui présentent un certain nombre de nœuds (ou sommets) reliés par un certain nombre d'arêtes, ou arcs (selon la représentation). Un graphe est un peu comme une carte sur laquelle les localités seraient les nœuds, et les routes seraient les arcs. Cette représentation diffère de l'arbre, sur lequel chaque chemin se termine sur un nœud feuille. Un graphe est représenté sur la Figure 3.2.

Les graphes sont particulièrement utiles lorsque les états représentent un espace physique. Le GPS, par exemple, utilise un graphe pour représenter les localités et les voies de circulation. Les graphes ajoutent aussi quelques subtilités auxquelles vous n'auriez peut-être pas songé. Ainsi, par exemple, un graphe peut inclure le concept de direction. Contrairement à l'arbre, dans lequel les relations sont de type père/fils, un nœud dans un graphe peut être relié à n'importe quel autre nœud et la relation peut être envisagée dans un sens particulier. Il en est de même que des rues dans une ville : elles sont le plus souvent bidirectionnelles, mais certaines sont à sens unique.

La représentation d'une relation dans un graphe ne reflète pas nécessairement les réalités du graphe. Un coefficient peut être assigné à une relation particulière. Ce coefficient peut représenter la distance entre deux points, ou la durée de la portion de chemin, ou une autre information quelconque.



Un arbre n'est rien d'autre qu'un graphe dans lequel deux sommets ne peuvent être reliés que par un seul chemin, et qui ne comporte pas de cycles (le cycle étant une relation permettant de revenir du fils au père). Souvent, les algorithmes de graphes ne s'appliquent qu'aux arborescences.

Traverser un graphe signifie parcourir (visiter) chaque sommet (nœud) dans un ordre particulier. Ce processus peut inclure la lecture et la mise à jour de l'information associée à chaque sommet. En traversant un graphe, on atteint des sommets non visités, qui deviennent des sommets visités (dès lors qu'on les a atteints) ou traités (lorsque l'algorithme a parcouru tous les arcs qui partent de ce sommet). L'ordre du parcours est lié au type de recherche, non informée (aveugle) ou informée (heuristique). Dans le cas de la *recherche non informée*, l'IA explore l'espace d'états sans autre information que la structure du graphe, qu'elle découvre à mesure qu'elle le

traverse. Les sections suivantes sont consacrées à deux algorithmes de recherche aveugle couramment utilisés : un algorithme de parcours en largeur et un algorithme de parcours en profondeur.

Un algorithme de parcours en largeur commence à la racine du graphe et explore chacun des nœuds qui lui sont reliés. Il poursuit le parcours au niveau suivant et explore tous les niveaux l'un après l'autre jusqu'à ce qu'il atteigne la fin du graphe. Ainsi, dans notre exemple, le parcours se fait de A à B puis de B à C, et enfin, de C à D. Le parcours du graphe s'effectue de façon systématique et circulaire. L'algorithme visite d'abord tous les sommets du premier niveau, c'est-à-dire ceux directement reliés au sommet de départ, puis il explore le deuxième niveau, puis le troisième, et ainsi de suite.

Un algorithme de parcours en profondeur commence à la racine du graphe et explore les nœuds successifs sur un chemin unique jusqu'au sommet final du graphe. Il revient ensuite par un deuxième chemin, non encore parcouru, jusqu'au sommet racine. Si d'autres chemins à partir de la racine sont possibles, il en choisit un et effectue à nouveau un parcours. Il s'agit d'explorer entièrement chaque chemin avant d'en explorer un autre.

Jouer à des jeux qui opposent les joueurs

Ce qui est intéressant dans les parcours d'espaces d'états, c'est qu'ils représentent à la fois les fonctionnalités actuelles de l'IA et ses possibilités futures. Il en est ainsi de tous les jeux dans lesquels des joueurs poursuivent un objectif qui est en conflit avec les objectifs des autres. Un jeu simple comme le morpion est l'exemple parfait d'un parcours d'espaces dans lequel vous avez peut-être déjà

vu l'IA intervenir. Dans le film *WarGames* (1983), le supercalculateur WOPR (*War Operation Plan Response*) joue contre lui-même à une vitesse impressionnante, mais ne peut pas gagner car le jeu est extrêmement simple et il suffit d'utiliser la technique du parcours d'espace d'états pour ne jamais perdre.

Au morpion, le jeu est constitué de neuf cellules qu'un des deux joueurs peut remplir avec des X, et l'autre joueur avec des O. Le joueur qui arrive, le premier, à aligner trois de ses propres éléments (sur une ligne horizontale, verticale ou diagonale) a gagné. Dans la construction d'une arborescence d'espaces d'états, chaque niveau de l'arborescence représente un tour. Les nœuds finals représentent l'état final du jeu et correspondent à une victoire, une égalité ou une défaite pour l'IA. À chaque nœud terminal correspond un score plus élevé pour gagner, plus faible pour l'égalité et plus faible encore, ou négatif, pour la défaite. L'IA assigne les scores aux nœuds supérieurs et aux branches par addition jusqu'à avoir atteint le nœud de départ. Le nœud de départ représente la situation courante. Une stratégie simple vous permet de traverser l'arborescence : quand vient le tour de l'IA et quand vous devez assigner les valeurs d'un certain nombre de nœuds, vous calculez la valeur maximum (l'IA devant sans doute obtenir le résultat maximum du jeu), et quand c'est son adversaire qui doit jouer, on calcule la valeur minimum. Au bout du compte, on obtient un arbre sur lequel des scores sont attribués aux branches. L'IA choisit son coup en fonction de la branche dont la valeur est la plus élevée, car cela correspond au parcours des nœuds qui offrent la meilleure possibilité de gagner. La Figure 3.3 illustre cette stratégie.

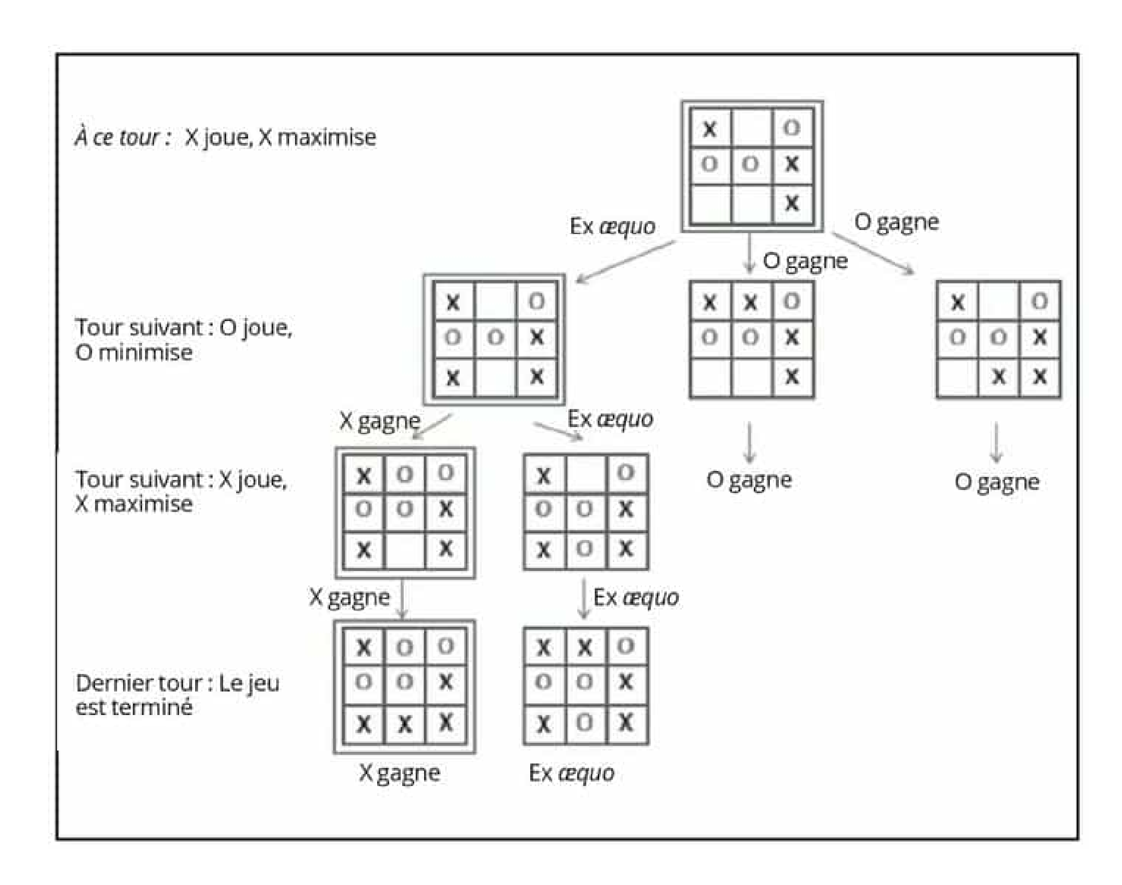


FIGURE 3.3 : L'approximation du minimax dans une partie de morpion.

Cette méthode est appelée l'approximation du minimax. Elle a été présentée en 1987 par Ronald L. Rivest, du laboratoire d'informatique du MIT (vous pouvez lire son article à l'adresse https://people.csail.mit.edu/rivest/pubs/Riv87c.pdf). Depuis, cet algorithme et ses variantes ont trouvé des applications dans divers jeux compétitifs, parallèlement à des avancées récentes comme AlphaGo, de DeepMind, qui utilise une méthode inspirée de l'approximation du minimax (qu'on retrouve aussi dans le film *WarGames*).



À propos de l'approximation du minimax, on entend parfois parler de l'élagage *alpha-bêta*, qui est une méthode subtile pour réduire le nombre de nœuds à évaluer dans les espaces d'états complexes. Dans certains jeux, les arborescences ne sont pas compactes

du tout : elles peuvent comporter des millions de branches, et il convient alors de procéder à un élagage pour limiter le nombre de calculs.

Avoir recours à une recherche locale et à une heuristique

La méthode du parcours d'espaces d'états recouvre beaucoup de choses. Au bout du compte, aucune machine, quelle que soit sa puissance, ne peut énumérer toutes les possibilités auxquelles une situation peut donner lieu. Cette section traite également des jeux, sachant qu'ils sont prédictibles et que leurs règles sont fixes, tandis que dans la vie réelle un grand nombre de situations sont imprédictibles et n'obéissent pas à des règles claires, si bien que les jeux constituent un contexte optimiste et privilégié.

Au jeu de dames, qui est relativement simple comparé au jeu d'échecs ou au jeu de Go, il existe tout de même 500 milliards de milliards (500000000000000000000) de positions possibles sur la grille, et selon des calculs effectués par les mathématiciens de l'université d'Hawaii, ce nombre serait égal au nombre de grains de sable sur la terre. Il est vrai que le nombre de coups possibles diminue à mesure que la partie progresse. Cependant, le nombre de possibilités à évaluer à chaque coup reste extrêmement élevé. Pour obtenir ce résultat de 500 milliards de milliards de possibilités, il a fallu 18 années de calcul en utilisant des calculateurs performants (http://sciencenetlinks.com/science-news/science-updates/checkers-solved/). Imaginez le temps qu'il faudrait avec votre ordinateur personnel pour calculer ne serait-ce qu'un sous-ensemble de ces possibilités de mouvement. Pour que ce soit

gérable, il serait nécessaire de se limiter à un très petit sous-ensemble.

L'optimisation à l'aide d'une recherche locale et d'une heuristique consiste à utiliser des contraintes pour limiter le nombre d'évaluations possibles au départ (comme pour l'élagage alpha, on renonce à une partie des calculs parce que ces calculs ne contribueraient en rien à la réalisation de l'objectif). La *recherche locale* est une méthode générale de résolution des problèmes faisant appel à un vaste ensemble d'algorithmes et permettant d'éviter les complexités exponentielles de nombreux problèmes de type NP. Une recherche locale prend pour point de départ une solution imparfaite, et s'en éloigne par étapes successives. Elle détermine la viabilité des solutions voisines et peut aboutir à la solution parfaite, en se fondant sur un choix aléatoire ou sur une heuristique astucieuse (aucune méthode exacte n'entre en jeu).



Une *heuristique* est une estimation éclairée concernant une solution, comme par exemple une règle simple pour se diriger vers un résultat désiré mais sans pouvoir dire exactement comment y parvenir. C'est comme être perdu dans une ville que l'on ne connaît pas et se faire indiquer un chemin pour retrouver l'hôtel (mais sans instructions précises), ou simplement la distance à laquelle on se trouve de l'hôtel.

Les algorithmes de recherche locale progressent par itérations à partir d'une situation de départ, étape par étape, en passant d'une solution à une autre solution meilleure dans l'espace d'états jusqu'à ce qu'il ne soit plus possible d'améliorer la solution. Parce qu'ils sont très simples et très intuitifs, il n'est pas difficile de concevoir une méthode de recherche locale pour résoudre un problème d'algorithmique. En général, rendre cette méthode performante est plus difficile. Tout dépend de la définition de la procédure adéquate :

1. **Commencer par une solution existante (généralement une solution aléatoire ou une solution donnée par un autre algorithme).**
2. **Rechercher une série de nouvelles solutions possibles au voisinage de la solution courante, constituant ainsi une liste de solutions candidates.**
3. **Déterminer laquelle de ces solutions doit être utilisée à la place de la solution courante, d'après le résultat d'une heuristique acceptant en entrée la liste des solutions candidates.**
4. **Poursuivre l'exécution des étapes 2 et 3 jusqu'à ce qu'il n'apparaisse plus aucune amélioration possible de la solution courante, ce qui signifie que l'on a abouti à la meilleure solution disponible.**

Bien qu'elle soit facile à concevoir, une recherche locale ne permet pas toujours de trouver une solution dans un temps raisonnable (on peut interrompre le processus et utiliser la solution courante) ni de produire une solution de qualité minimale. Il n'est pas garanti qu'une recherche locale aboutisse à la solution du problème, mais les chances s'améliorent depuis le point de départ quand on dispose d'un temps suffisant pour l'exécution des calculs. Le processus ne s'arrête qu'une fois qu'il ne trouve plus aucun moyen d'améliorer la solution. Le secret consiste à déterminer le voisinage qu'il convient d'explorer. Si vous explorez tout, ce sera finalement une recherche exhaustive, ce qui implique une explosion du nombre de possibilités à explorer et à tester.

Le recours à l'heuristique limite le champ d'investigation sur la base d'une règle pratique. L'heuristique est parfois aléatoire, et si la méthode est « non intelligente », la solution peut néanmoins faire l'affaire. Peu de gens

savent, par exemple, qu'à l'origine, l'aspirateur robot autonome Roomba, conçu par trois diplômés du MIT, se déplaçait de façon purement aléatoire. Il était pourtant considéré comme un appareil intelligent et faisait un excellent travail (ici, l'intelligence réside en réalité dans l'idée de recourir à une solution aléatoire pour résoudre un problème qui, autrement, était trop compliqué).

Le choix aléatoire n'est pas la seule heuristique possible. Une recherche locale peut reposer sur des solutions d'exploration plus raisonnées, grâce à des heuristiques bien conçues pour définir des directions, par exemple avec la méthode du *hill-climbing* ou l'algorithme *Twiddle*, en évitant le piège qui consiste à accepter des solutions médiocres, comme avec le *recuit simulé* et la *recherche tabou*. Le *hill-climbing*, le *Twiddle*, le *recuit simulé* et la *recherche tabou* sont des algorithmes de recherche qui utilisent efficacement des heuristiques.

La méthode du *hill-climbing*, dite aussi de l'*escalade*, s'inspire de la force de gravitation. Elle repose sur l'idée qu'un ballon qui dévale une pente suit le chemin le plus pentu. Quand il gravit un monticule, le ballon a tendance à suivre le chemin le plus direct vers le sommet, qui est un des chemins les plus escarpés. Le problème de l'IA est analogue, et l'heuristique est une règle indiquant la meilleure méthode pour descendre ou monter, parmi les états possibles de l'espace des états. C'est un algorithme efficace, mais avec parfois des blocages, dans des situations appelées plateaux (creux intermédiaires) ou pics (maxima locaux).

L'algorithme Twiddle, ou algorithme de descente par coordonnées, est similaire aux algorithmes d'escalade. L'heuristique consiste à explorer toutes les directions possibles, mais en concentrant la recherche dans la direction du voisinage le plus efficient. L'algorithme calibre les étapes de sa progression, en ralentissant quand il lui

est difficile de trouver de meilleures solutions, jusqu'à ce qu'il atteigne un point d'arrêt.

Le *recuit simulé* tire son nom d'une technique utilisée en métallurgie, le recuit, consistant à chauffer le métal à haute température, puis à le laisser refroidir lentement afin de l'assouplir pour le travailler à froid et éliminer les défauts internes de cristallisation. La recherche locale reprend ce principe en assimilant la recherche de solution à une structure atomique qui évoluerait de manière à devenir plus maniable. La température est l'élément qui change la donne dans le processus d'optimisation. De même que les hautes températures assouplissent la structure du matériau (les solides fondent et les liquides s'évaporent), les hautes températures dans un algorithme de recherche locale entraînent un assouplissement de la fonction objectif, si bien que les moins bonnes solutions peuvent être préférées aux meilleures. Le recuit simulé modifie la procédure d'escalade en réservant la fonction objectif pour l'évaluation des solutions voisines, mais en l'utilisant pour déterminer le choix d'une façon différente.

La recherche tabou recourt à la mémorisation pour déterminer les parties du voisinage à explorer. Quand une solution semble avoir été trouvée, l'algorithme retrace d'autres chemins possibles qu'il n'avait pas encore explorés afin de vérifier quelle est la meilleure solution.

Le recours à des mesures de direction (ascendante ou descendante) ou de température (contrôle du processus aléatoire), la simple restriction ou la reconstitution d'une partie de la recherche, sont autant de moyens d'éviter de devoir tout essayer et de se concentrer sur une bonne solution. Prenons l'exemple d'un robot qui marche. On guide le robot dans un environnement inconnu pour chercher à atteindre une cible particulière tout en évitant les obstacles. Dans le domaine de l'intelligence artificielle, c'est à la fois une tâche fondamentale et un défi de taille. Dans sa navigation, le robot peut être guidé par un

télémètre laser (lidar, *light detection and ranging*) ou par un sonar (un dispositif permettant, pour ainsi dire, de visualiser l'environnement grâce au son). Cependant, quel que soit le degré de sophistication du matériel, il faut des algorithmes adéquats pour exécuter les tâches suivantes :

» Trouver le plus court chemin vers une destination (ou du moins, un chemin raisonnablement court).

» Éviter les obstacles sur le chemin.

» Manifester des comportements personnalisés, comme tourner ou freiner le moins possible.

Un algorithme de recherche de chemin permet au robot, à partir d'un point de départ donné, d'atteindre un objectif en utilisant le plus court chemin parmi deux, tout en anticipant et en évitant les obstacles le long de sa trajectoire (il ne suffit pas qu'il réagisse après avoir heurté un mur). La recherche de chemin est utile également pour déplacer un appareil vers une cible dans un espace, lequel peut même être virtuel, par exemple dans le cas d'un jeu vidéo ou de pages Internet. En l'occurrence, le robot perçoit le mouvement comme un flux d'espaces d'états vers les frontières de ses capteurs. Si l'objectif est en dehors du champ, le robot ne saura pas où aller. L'heuristique peut l'orienter dans la bonne direction (il saura, par exemple, que la cible se trouve au nord) et lui permettre d'éviter les obstacles dans des temps acceptables sans qu'il soit nécessaire de déterminer tous les chemins possibles pour cela.

Découvrir la machine qui apprend

Tous les exemples d'algorithmes abordés jusqu'ici sont associés à l'IA, sachant que ce sont des solutions

intelligentes pour résoudre des problèmes répétitifs et bien circonscrits, bien que complexes, pour lesquels une intelligence est requise. Il faut qu'un architecte étudie le problème et choisisse le bon algorithme pour le résoudre. Les changements, les mutations et la manifestation de caractéristiques inhabituelles peuvent devenir un réel problème lorsqu'il s'agit d'obtenir une exécution probante de l'algorithme. En effet, l'apprentissage du problème et de sa solution est un processus qui se produit une fois pour toutes, au moment où l'algorithme intervient dans le logiciel. Vous pouvez, par exemple, programmer une IA pour résoudre une grille de Sudoku (un jeu à la mode consistant à inscrire des chiffres dans une grille en respectant certaines règles : https://www.sudokugratuit.com/comment-jouer.html). Vous pouvez même rendre l'algorithme assez flexible pour qu'il accepte, par la suite, davantage de règles, ou bien des grilles plus grandes. Peter Norvig, directeur de recherche chez Google, a rédigé sur ce sujet un essai extrêmement intéressant (http://norvig.com/sudoku.html) qui montre comment utiliser judicieusement la recherche en profondeur et limiter le nombre de calculs (sans quoi les calculs pourraient durer un temps infini), utiliser des contraintes et explorer d'abord les branches les plus courtes pour trouver des solutions au Sudoku.

Malheureusement, tous les problèmes ne peuvent pas se résoudre de la même manière qu'une grille de Sudoku. Les problèmes réels ne sont jamais formulés par des mots simples, sous forme d'informations complètes et avec une action bien définie. Voyons ce qui caractérise, par exemple, le problème de la détection d'une escroquerie aux assurances, ou bien celui du diagnostic d'une pathologie :

» **Un vaste ensemble de règles et de possibilités :** le nombre de fraudes possibles est extrêmement

élevé ; plusieurs pathologies différentes peuvent faire apparaître les mêmes symptômes.

- **Des informations manquantes :** les fraudeurs peuvent dissimuler certaines informations ; les médecins doivent souvent se contenter d'informations incomplètes (il se peut qu'un examen manque).
- **Les règles ne sont pas immuables :** les fraudeurs peuvent découvrir de nouveaux moyens de frauder ou d'arnaquer ; on voit apparaître ou on découvre de nouvelles pathologies.

Pour résoudre de tels problèmes, il n'est pas possible de se contenter d'une méthode prédéterminée. Une approche flexible s'impose, et il est nécessaire d'accumuler des connaissances pratiques pour pouvoir faire face à un nouveau défi. En d'autres termes, il s'agit de continuer d'apprendre, comme nous devons le faire tout au long de notre existence afin d'être capables de faire face à des conditions changeantes et difficiles.

Tirer parti des systèmes experts

Les systèmes experts ont été la première tentative d'échapper aux contraintes qu'imposaient les algorithmes codés en dur et de trouver des méthodes plus flexibles et plus subtiles pour résoudre des problèmes réels. L'idée de base était simple et pertinente, à une époque où le stockage en mémoire et le traitement de grandes quantités de données restaient onéreux. Cela peut sembler étrange aujourd'hui, mais dans les années 1970 un spécialiste de l'IA comme Ross Quinlan avait été obligé de montrer comment construire des modèles de langage viables basés sur un vocabulaire

de 20 mots seulement, car la mémoire d'un ordinateur ne pouvait pas supporter plus. Les possibilités étaient limitées, et une solution consistait à cibler les informations essentielles et à les obtenir des personnes qui maîtrisaient le mieux le problème.



Les systèmes experts étaient experts non pas parce que leurs connaissances étaient fondées sur leur propre processus d'apprentissage, mais plutôt parce qu'ils les obtenaient d'experts humains qui fournissaient un système prédigéré d'informations essentielles tirées de l'étude d'ouvrages, apprises d'autres spécialistes ou découvertes par eux-mêmes. Fondamentalement, c'était une façon intelligente de transférer la connaissance vers une machine.

Un des premiers systèmes de ce genre est MYCIN, conçu pour diagnostiquer les maladies liées à la coagulation du sang et des infections comme la bactériémie (infection du sang par une bactérie) et la méningite (inflammation des membranes qui protègent le cerveau et la moelle épinière). MYCIN recommandait le dosage adéquat d'antibiotiques en appliquant plus de 500 règles et en demandant confirmation, le cas échéant, au médecin qui l'utilisait. Quand les informations n'étaient pas suffisantes, par exemple s'il manquait des résultats d'examens de laboratoire, MYCIN entamait un dialogue consultatif en posant les bonnes questions afin de parvenir à un diagnostic fiable et à la prescription d'une thérapie adaptée.

Écrit en Lisp par Edward Shortliffe dans le cadre d'une thèse de doctorat à l'université Stanford, MYCIN a nécessité plus de cinq années de travail et s'est révélé plus performant que n'importe quel médecin assistant ou interne. En effet, ses diagnostics étaient aussi pertinents et précis que ceux d'un médecin expérimenté. C'est dans le même laboratoire qu'avait été conçu DENDRAL, le premier système expert jamais créé, quelques années

plus tôt. DENDRAL est une application exigeante, spécialisée en chimie organique, pour laquelle les algorithmes de force brute se sont révélés inutilisables dans le contexte d’une heuristique humaine fondée sur une expérience de terrain.

Le succès de MYCIN n’a pas été sans poser certains problèmes. Tout d’abord, la définition des responsabilités n’était pas claire (au cas où le système produirait un diagnostic incorrect, qui serait responsable ?). Ensuite, l’utilisation de MYCIN posait un problème dans la mesure où le médecin devait se connecter au grand système de Stanford à partir d’un terminal distant, une procédure difficile et lente en un temps où l’Internet n’en était qu’à ses balbutiements. MYCIN a cependant fait la preuve de son efficacité et de son utilité dans l’aide à la décision humaine, et il a ouvert la voie aux nombreux autres systèmes experts qui se sont multipliés par la suite, dans les années 1970 et 1980.

À cette époque, les systèmes experts étaient généralement constitués de deux éléments bien distincts : une base de connaissances et un moteur d’inférence. La base de connaissances assimilait la connaissance à un ensemble de règles sous forme de déclarations de type si… alors (l’instruction « si » étant associée à une ou plusieurs conditions et l’instruction « alors » à des énoncés conclusifs). Ces déclarations revêtaient une forme symbolique, avec une différenciation entre les occurrences (épreuves ou faits simples), les classes et les sous-classes, qui pouvaient toutes être manipulées en utilisant une logique booléenne ou une logique du premier ordre élaborée, avec davantage d’opérations possibles.



La *logique du premier ordre* est un ensemble d’opérations qui ne consistent pas simplement à combiner des assertions vraies (VRAI) et fausses (FAUX). Elle fait notamment intervenir des concepts comme POUR TOUT et IL EXISTE, ce qui permet de gérer des propositions qui

peuvent être vraies mais sans pouvoir être prouvées par les éléments dont on dispose sur le moment. Pour en savoir plus sur cette forme de logique, vous pouvez tout d'abord lire l'article suivant : http://whatis.techtarget.com/definition/first-order-logic.

Le *moteur d'inférence* est une série d'instructions qui disent au système comment manipuler les conditions sur la base d'un ensemble logique booléen d'opérateurs comme AND, OR, NOT. Les conditions symboliques VRAI (une règle s'applique) et FAUX (la règle ne s'applique pas) peuvent être combinées pour former un raisonnement complexe.

Le système ayant été fondé sur une série de « si » (conditions) et de « alors » (conclusions), et structuré en couches successives, l'acquisition d'informations initiales a permis d'écarter certaines conclusions tout en permettant aussi au système d'interagir avec l'utilisateur, concernant les informations pouvant conduire à une réponse. Quand on travaillait avec le moteur d'inférence, les opérations couramment effectuées par les systèmes experts étaient les suivantes :

» **Chaînage avant :** les éléments disponibles déclenchaient une série de règles et en excluaient d'autres à chaque étape. Le système fonctionnait initialement avec des règles qui pouvaient déclencher de nouvelles inférences. Cette approche est clairement pilotée par les données.

» **Chaînage arrière :** le système évalue toutes les conclusions possibles et les évalue en fonction des données disponibles. Cette méthode pilotée par les objectifs permet de déterminer les questions à poser et élimine des séries entières d'objectifs. MYCIN utilisait le chaînage arrière ; progresser à l'envers, de l'hypothèse vers les données, est une stratégie

couramment employée pour les diagnostics médicaux.

» **Résolution de conflits** : quand un système aboutit à plus d’une conclusion à la fois, il favorise la conclusion qui présente certaines caractéristiques (en termes d’impact, de risque ou d’autres facteurs). Parfois, le système consulte l’utilisateur, et la résolution est effectuée sur la base des évaluations faites par ce dernier. MYCIN, par exemple, utilisait un facteur de certitude qui estimait la probabilité d’exactitude du diagnostic.

Un grand avantage de ces systèmes était qu’ils représentaient la connaissance sous une forme lisible par l’homme, si bien que la décision était transparente et pouvait être comprise et modifiée. Quand le système aboutit à une conclusion, il retourne les règles utilisées pour y aboutir. L’utilisateur peut vérifier le travail du système de façon systématique et l’accepter, ou identifier des signes d’erreur d’entrée et procéder aux corrections. Par ailleurs, les systèmes experts étaient faciles à programmer à l’aide de langages comme Lisp ou l’ALGOL. Avec le temps, les utilisateurs amélioraient les systèmes experts en ajoutant de nouvelles règles ou en modifiant les règles existantes. Il était même possible de les faire fonctionner dans des conditions d’incertitude en appliquant une logique floue (une logique avec laquelle une valeur peut contenir n’importe quel nombre compris entre 0, qui signifie absolument faux, et 1, qui signifie absolument vrai). La logique floue permet d’éviter les changements abrupts liés au déclenchement d’une règle à partir d’un seuil. Par exemple, si une règle doit se déclencher quand il fait chaud dans la pièce, plutôt que d’être déclenchée pour une température exacte, elle sera déclenchée lorsque la température est proche de ce seuil. Si les systèmes experts ont connu le déclin à la fin des

années 1980 et si leur développement s'est arrêté, c'est principalement pour les raisons suivantes :

> » La logique et le symbolisme de ces systèmes ont révélé leurs limites pour ce qui est d'exprimer les règles sur lesquelles se fonde une décision, ce qui a conduit à la création de systèmes personnalisés, c'est-à-dire au retour à des règles codées en dur avec des algorithmes classiques.
>
> » Pour de nombreux problèmes difficiles, les systèmes experts devenaient si complexes et délicats qu'ils perdaient leur attrait en termes de faisabilité et de coût économique.
>
> » Les données devenant plus répandues et plus disponibles, il n'était plus justifié de se donner la peine d'aller recueillir et distiller des connaissances spécialisées alors que ces mêmes connaissances (voire des connaissances meilleures encore) pouvaient être extraites des données.

Les systèmes experts existent toujours. Ils sont utilisés dans l'évaluation du crédit et la détection des fraudes, et dans d'autres domaines encore, sachant qu'ils ne doivent pas simplement fournir une réponse, mais aussi établir de façon claire et transparente les règles sur lesquelles se fonde une décision, d'une manière qui soit acceptable du point de vue de l'utilisateur du système (comme du point de vue d'un spécialiste du domaine en question).

Faire intervenir l'apprentissage machine

Des solutions pour l'apprentissage direct à partir des données, sans qu'il soit nécessaire de les formater au préalable, ont vu le jour plusieurs décennies avant les

systèmes experts. Certaines relevaient de la statistique, tandis que d'autres consistaient à imiter la nature de différentes manières, et d'autres encore consistaient à produire de façon autonome une logique symbolique sous forme de règles à appliquer à l'information brute. Toutes ces solutions provenaient de différentes écoles et portaient des noms différents. Aujourd'hui, on parle d'apprentissage machine. L'apprentissage machine appartient au domaine des algorithmes, mais contrairement aux algorithmes dont il a été question jusqu'ici, il ne consiste pas à résoudre un problème au moyen d'une série d'étapes prédéfinies. En règle générale, l'apprentissage machine concerne des problèmes que nous ne savons pas définir sous forme d'un processus par étapes, mais que nous résolvons naturellement. C'est le cas, par exemple, de la reconnaissance des visages sur des photos et de l'identification de certains mots dans une conversation. L'apprentissage machine est mentionné dans tous les chapitres de ce livre, mais les Chapitres 9 et 11 expliquent le fonctionnement des principaux algorithmes, en particulier ceux de l'apprentissage profond, la technologie à la base de la nouvelle série d'applications de l'IA dont la presse parle presque tous les jours.

Atteindre de nouveaux sommets

Le rôle de l'apprentissage machine dans la nouvelle série d'algorithmes d'IA est en partie de remplacer et en partie de compléter les algorithmes existants, en rendant accessibles des activités qui requièrent une intelligence d'un point de vue humain, laquelle intelligence ne peut pas facilement être mise sous forme d'une séquence précise d'étapes. Un exemple évident de ce rôle est celui de la maîtrise dont fait preuve le champion du jeu de Go qui, en

un coup d'œil, apprécie les menaces et les possibilités que présente une configuration, puis joue en conséquence (pour un historique du jeu de Go, voir http://jeudego.org/_php/histoiredugo.php). Le Go est un jeu extrêmement complexe pour une IA. Alors qu'aux échecs une configuration donnée permet d'envisager en moyenne 35 mouvements possibles et qu'une partie comporte souvent plus de 80 mouvements, le Go présente 140 mouvements à évaluer et une partie de Go comporte généralement plus de 240 mouvements. Actuellement, aucune puissance de calcul au monde ne permet de créer un espace d'états complet pour une partie de Go. À Londres, l'équipe de DeepMind, aujourd'hui filiale de Google, a développé AlphaGo, un logiciel qui a vaincu un certain nombre de joueurs classés (voir https://deepmind.com/research/alphago/). Ce logiciel n'est pas fondé sur une méthode algorithmique d'exploration d'un immense espace d'états, mais sur les processus suivants :

- Une méthode de recherche intelligente fondée sur des tests aléatoires des mouvements possibles. L'IA applique de façon répétitive un parcours en profondeur pour déterminer si le premier résultat trouvé est positif ou négatif (un espace d'états incomplet et partiel).
- Un algorithme d'apprentissage profond qui traite une image du jeu (un instantané) et qui en tire à la fois le meilleur mouvement possible (l'algorithme est appelé le *réseau d'objectifs*) et une estimation de la probabilité que l'IA gagne la partie grâce à ce mouvement (l'algorithme est appelé le *réseau de valeur*).
- Une capacité d'apprendre des parties déjà jouées par les champions du Go et par l'IA contre elle-même, comme WOPR dans le film *WarGames* (1983). Une version récente de ce programme,

> AlphaGo Zero, est capable d’apprendre seule, sans aucun exemple humain. Cette capacité d’apprentissage est ce que l’on appelle l’*apprentissage par renforcement*.

Chapitre 4

Innover avec un matériel spécialisé

DANS CE CHAPITRE…

- **Utiliser un matériel informatique standard**
- **Utiliser un matériel informatique spécialisé**
- **Améliorer le matériel**
- **Interagir avec l’environnement**

Au Chapitre 1, nous avons vu qu’une des raisons de l’échec des premiers projets d’IA était l’absence de matériel informatique adéquat. Les ordinateurs de l’époque n’étaient pas capables d’exécuter assez rapidement les tâches, même pour des besoins banals, et à plus forte raison pour une entreprise aussi complexe que l’imitation de la pensée humaine. Ce sujet est abordé en détail dans le film *Imitation Game* (https://www.amazon.com/exec/obidos/ASIN/B00RY86HSU/datacservip0f-20/), qui montre comment Alan Turing a finalement craqué le code Enigma en ayant l’idée subtile de recherche l’expression « Heil Hitler » dans chaque message. Sans cette faille particulière dans le mode d’utilisation de la machine Enigma, le matériel de calcul utilisé par Turing n’aurait jamais été assez rapide pour pouvoir résoudre le problème (et il n’y aurait pas eu matière à réaliser un film). Les témoignages historiques –

le peu qui est déclassifié – montrent que les problèmes rencontrés par Turing ont été plus complexes que ce que le film laisse paraître. Aujourd'hui, heureusement, les machines les plus courantes offrent une vitesse de traitement suffisante pour résoudre un grand nombre de problèmes, et c'est l'objet du début de ce chapitre.



Pour pouvoir vraiment commencer à imiter la pensée humaine, il faut un matériel spécialisé, et pour l'instant, même le meilleur ordinateur spécialisé n'est pas adapté à cette tâche. Pratiquement tous les matériels habituels sont conçus selon l'architecture de von Neumann (http://www.c-jump.com/CIS77/CPU/VonNeumann/lecture.html), qui sépare la mémoire du calcul et crée ainsi un système de traitement formidablement généraliste, qui n'est tout simplement pas adapté à certains types d'algorithmes car la vitesse de transmission du bus entre le processeur (ou CPU, pour *central processing unit*) et la mémoire n'est pas suffisante pour éviter un *goulet d'étranglement de von Neumann*. La seconde partie de ce chapitre vous fait découvrir les différentes méthodes utilisées pour résoudre le problème posé par le goulet d'étranglement de von Neumann et accélérer l'exécution des algorithmes complexes qui doivent gérer des données en grand nombre.

Même un matériel informatique spécialement conçu pour exécuter des calculs rapidement et imiter la pensée humaine ne peut pas fonctionner plus vite que ce que lui permettent ses entrées et ses sorties. C'est pourquoi des recherches sont menées pour créer un environnement dans lequel le matériel fonctionnerait mieux. Le problème peut être résolu de plusieurs manières, mais ce chapitre en examine deux : accroître les capacités du matériel informatique, et utiliser des capteurs spéciaux. Ces changements apportés à l'environnement matériel sont utiles, mais comme nous l'expliquons plus loin, ils ne

sauraient suffire pour fabriquer l'équivalent d'un cerveau humain.

En fin de compte, quels que soient les progrès qui lui sont apportés, l'ordinateur ne sert à rien si les humains qui l'utilisent ne sont pas capables d'interactions efficaces avec la machine. La dernière section de ce chapitre décrit des techniques pour améliorer ces interactions, qui sont simplement le résultat d'un progrès dans la production des extrants et d'une programmation subtile. De même qu'Alan Turing avait trouvé un truc pour que son ordinateur paraisse plus capable qu'il ne l'était réellement, ces techniques font que les ordinateurs semblent accomplir de véritables miracles. En réalité, l'ordinateur ne comprend rien : tout le mérite revient à ceux qui le programment.

Utiliser un matériel standard

La plupart du temps, pour commencer un projet d'IA, vous utiliserez tout d'abord un matériel standard, car les composants actuellement disponibles dans le commerce vous offrent une puissance de traitement considérable, surtout comparés aux composants des années 1980, l'époque où l'IA avait commencé à produire des résultats exploitables. Par conséquent, même si ce matériel standard ne permet finalement pas le même niveau de production qu'un matériel spécialisé, il est possible d'aller suffisamment loin avec un code expérimental pour créer un modèle viable qui traitera un jeu complet de données.

Comprendre en quoi consiste un matériel standard

L'*architecture* (structure) du PC standard n'a pas changé depuis le jour où John von Neumann en avait fait la promotion pour la première fois en 1946 (lire l'article à l'adresse https://www.maa.org/external_archive/devlin/devlin_12_03.html). L'historique de la page https://lennartb.home.xs4all.nl/coreboot/col2.html montre que dans les ordinateurs individuels, le processeur est relié à la mémoire et aux périphériques par un bus depuis 1981 (et même, depuis bien avant). Tous ces systèmes utilisent l'architecture de von Neumann parce que cette architecture présente d'importants avantages en termes de modularité. On peut lire que sur ces appareils, il est possible de décider de mettre à jour tel ou tel composant et que cela permet d'augmenter leur capacité. Sur n'importe quel PC, par exemple, vous pouvez, dans certaines limites, augmenter la quantité de mémoire ou l'espace de stockage disponible. Vous pouvez aussi utiliser des périphériques avancés. Néanmoins, tous ces éléments seront reliés par un bus.



La possibilité d'accroître les capacités d'un ordinateur personnel ne change rien à ce qui caractérise son architecture. Le PC que vous utilisez aujourd'hui est donc conçu selon la même architecture que des machines conçues il y a longtemps déjà : il est simplement plus performant. Par ailleurs, le facteur de forme d'un appareil n'a pas non plus d'incidence sur son architecture. Dans les voitures, la connectivité des ordinateurs de bord est assurée par un système de bus basé sur l'architecture de von Neumann (même si ce n'est pas le même type de bus, l'architecture est la même). Si vous ne pensez pas que tout cela soit sans incidence sur la conception d'un appareil, consultez le schéma fonctionnel d'un BlackBerry : http://mobilesaudi.blogspot.com/2011/10/all-blackberry-schematic-complete.html. Ce

schéma, là encore, est conçu selon une configuration de von Neumann. Par conséquent, pratiquement tout appareil que l'on peut concevoir aujourd'hui aura une architecture similaire, même si les facteurs de forme, les types de bus et les capacités essentielles varient.

Recenser les défauts d'un matériel standard

La capacité de créer un système modulaire apporte d'importants avantages, surtout pour les entreprises. La possibilité de retirer et de remplacer les composants un par un permet de limiter les coûts tout en permettant des améliorations progressives en termes de vitesse d'exécution et d'efficience. Cependant, ici comme ailleurs, tout avantage a ses revers. La modularité qu'apporte l'architecture de von Neumann s'accompagne de graves déficiences :

EN QUOI CONSISTE L’ARCHITECTURE DE TYPE HARVARD ?

Aujourd’hui, certains systèmes utilisent l’architecture de type Harvard afin d’accélérer les traitements. L’architecture de type Harvard, comme l’architecture de von Neumann, repose sur une topologie en bus. Cependant, avec l’architecture de von Neumann, l’ordinateur utilise un bus unique et un espace mémoire unique pour véhiculer et stocker les instructions et les données, tandis que dans une architecture de type Harvard, l’accès aux instructions et l’accès aux données se font par deux bus distincts et les zones de mémoire physique peuvent aussi être distinctes (pour une comparaison, voir `http://infocenter.arm.com/help/topic/com.arm.doc.faqs/ka3839.html`). L’utilisation de bus séparés permet au système de récupérer l’instruction suivante, tout en attendant que l’instruction en cours récupère les données de la mémoire. L’architecture de type Harvard est donc à la fois plus rapide et plus efficiente. Néanmoins, la fiabilité en souffre, car il y a alors deux points de défaillance pour chaque opération : le bus d’instructions et le bus de données.

Les microcontrôleurs, par exemple ceux qui font fonctionner votre four à micro-ondes, utilisent souvent une architecture de type Harvard. Cette architecture est aussi utilisée dans certaines applications inhabituelles, pour une raison particulière. L’iPhone et la Xbox 360 utilisent des versions modifiées de l’architecture de type Harvard, dans lesquelles il existe une seule zone de stockage (au lieu de deux), mais des bus séparés. Ici, ce type d’architecture est utilisé pour la gestion numérique des droits (GND). Le code peut être paramétré en lecture seule pour que personne ne puisse le modifier ni créer de nouvelles applications sans autorisation. Du point de vue de l’IA, cela peut poser un problème, car une des propriétés de l’IA est précisément la capacité d’écrire de nouveaux algorithmes (du code exécutable) dans la mesure du nécessaire pour faire

face à des situations imprévues. Sachant que les ordinateurs individuels utilisent rarement l'architecture de type Harvard, que ce soit dans sa forme pure ou du point de vue de l'organisation en bus, ce livre ne la traite pas en détail.

- **Le goulet d'étranglement de von Neumann :** il s'agit de la plus grave de toutes ces déficiences, si l'on considère les exigences de disciplines comme l'IA, l'apprentissage machine ou même la science des données. Ce défaut particulier est étudié plus en détail dans la section « Prendre en compte le goulet d'étranglement de von Neumann », plus loin dans ce chapitre.
- **Des points de défaillance :** toute perte de connectivité au niveau du bus implique nécessairement une panne immédiate et brutale de l'ordinateur. Même dans les systèmes équipés de plusieurs processeurs, la cessation d'activité d'un seul processeur, au lieu d'entraîner une simple diminution de la capacité de traitement, provoque une panne totale. Le même problème se pose en cas de panne d'un autre composant du système : au lieu d'une limitation des fonctionnalités, c'est tout le système qui tombe en panne. Sachant que l'IA exige souvent le fonctionnement ininterrompu du système, le risque de conséquences graves s'accroît selon la manière dont une application est dépendante du matériel.
- **Une activité restreinte :** le bus de von Neumann peut soit récupérer une instruction, soit récupérer les données requises pour exécuter cette instruction, mais il ne peut pas faire l'un et l'autre au même moment. Par conséquent, lorsque l'extraction des données nécessite plusieurs cycles de bus, le processeur reste inutilisé, ce qui réduit d'autant plus

sa capacité d'exécution des tâches d'IA comportant de nombreuses instructions.

» **Une conception monotâche :** quand le cerveau entreprend une tâche, un certain nombre de synapses s'activent en même temps, ce qui permet l'exécution de plusieurs opérations. Or, le schéma initial de von Neumann ne permettait que l'exécution d'une seule opération à la fois, et seulement une fois que le système avait récupéré et l'instruction, et les données. Aujourd'hui, les ordinateurs sont généralement équipés de plusieurs processeurs capables d'exécuter des opérations de façon simultanée. Cependant, il faut encore que le code de l'application réponde à une telle exigence. De fait, cette fonctionnalité reste souvent inexploitée.

Utiliser des processeurs graphiques

Après avoir créé un prototype de configuration destiné à exécuter les tâches requises pour imiter la pensée humaine dans un domaine donné, vous aurez peut-être besoin d'un matériel supplémentaire afin de disposer d'une puissance de traitement qui vous permette d'exploiter l'ensemble des données exigées d'un système de production. Pour cela, il existe différentes possibilités mais une solution courante consiste à adjoindre des processeurs graphiques (des GPU) au processeur central de la machine. Les sections suivantes expliquent quelles sont les situations dans lesquelles un GPU est utile, ce que signifie précisément le terme GPU et pourquoi un GPU accélère le traitement.

Prendre en compte le goulet d’étranglement de von Neumann

Le goulet d’étranglement de von Neumann est ce qui résulte naturellement de l’utilisation d’un bus unique pour transférer les données entre le processeur, la mémoire vive, la mémoire de longue durée et les périphériques. Quelle que soit la vitesse à laquelle le bus accomplit sa tâche, il est toujours possible de le saturer, c’est-à-dire de provoquer un goulet d’étranglement ralentissant la transmission des données. Avec le temps, la vitesse des processeurs continue d’augmenter, tandis que les avancées qui concernent la mémoire et les autres éléments d’un système informatique portent plutôt sur la densité, c’est-à-dire la capacité de stocker davantage sur un espace plus réduit. En conséquence, le problème du goulet d’étranglement se pose avec plus d’acuité encore à chaque fois qu’un progrès est réalisé, et le processeur reste inactif pendant des laps de temps importants.

À PROPOS DE LA BOMBE DE TURING

La « Bombe » d'Alan Turing n'avait rien à voir avec l'IA. Il ne s'agissait pas non plus vraiment d'un ordinateur. C'était tout simplement une machine qui décryptait les messages envoyés par Enigma. Cette création a cependant été pour Turing l'occasion d'une réflexion approfondie, à la suite de laquelle il a finalement publié dans les années 1950 un article intitulé « Computing Machinery and Intelligence » (*https://www.abelard.org/turpap/turpap.php*). Dans cet article, Turing décrivait le *jeu de l'imitation*. La « Bombe » elle-même était en réalité inspirée d'une machine appelée Bomba et conçue par des Polonais.

Bien que certaines sources indiquent qu'Alan Turing travaillait seul, la Bombe a été fabriquée avec l'aide d'un grand nombre de personnes, et plus particulièrement celle de Gordon Welchman. Turing n'a pas non plus eu un beau jour l'illumination pour casser le code utilisé par les Allemands. À Princeton, il avait côtoyé des personnalités remarquables comme Albert Einstein et John von Neumann (qui allait par la suite inventer le concept de logiciel). Les articles rédigés par Turing ont inspiré d'autres scientifiques dans leur recherche de nouvelles possibilités.

Toutes sortes de matériels spécialisés continueront de voir le jour, aussi longtemps que des scientifiques continueront d'écrire des articles, d'échanger et de trouver de nouvelles idées et d'expérimenter. En regardant des films et en consultant divers médias, dont le contenu n'est pas nécessairement conforme aux faits historiques, ne vous imaginez pas que ces inventeurs sont des gens qui, un beau matin, se sont dit : « Aujourd'hui, je vais faire quelque chose de grandiose ! » Toute avancée se fonde sur quelque chose qui l'a précédée. Il est donc important d'étudier l'histoire des découvertes, car cela permet de cerner le chemin suivi et de mettre en lumière d'autres chemins prometteurs : ceux qui n'ont pas encore été suivis.

On peut raisonnablement surmonter une partie des problèmes qui sont liés au goulet d'étranglement de von Neumann et réaliser des progrès limités mais non négligeables en matière de rapidité des logiciels. Les solutions les plus courantes sont les suivantes :

» **L'antémémoire :** quand il est devenu évident que l'architecture de von Neumann ne permettait pas toujours d'extraire les données de la mémoire avec une rapidité suffisante, les fabricants d'ordinateurs ont vite réagi en ajoutant localement une mémoire ne nécessitant pas d'accès par bus. Cette mémoire apparaît comme externe au processeur, mais elle fait partie de son « package ». Cependant, sachant que l'antémémoire (ou cache) à haute performance coûte cher, les caches sont généralement de taille réduite.

» **Le cache de processeur :** malheureusement, les caches externes ne permettent pas encore une vitesse de traitement suffisante. Même en utilisant la RAM la plus rapide du marché et en court-circuitant entièrement l'accès par bus, on ne satisfait pas aux besoins en capacité de traitement du processeur. C'est pourquoi les fabricants se sont mis à ajouter de la mémoire interne : une antémémoire plus petite que le cache externe, mais avec un accès encore plus rapide sachant qu'elle est intégrée au processeur.

» **La préextraction :** le problème avec les antémémoires est qu'elles ne se révèlent utiles que lorsqu'elles contiennent les données correctes. Malheureusement, la mise en cache se révèle insuffisante dans les applications qui utilisent une grande quantité de données et exécutent un vaste ensemble de tâches. Pour rendre les processeurs plus rapides, l'étape suivante consiste à prévoir les données dont l'application devrait avoir besoin

l'instant d'après et à les charger en cache avant qu'elle ne les requière.

» **Une RAM spéciale :** il existe davantage de RAM différentes qu'on se l'imagine généralement, et il y a de quoi s'y perdre. Chaque type de RAM sert à résoudre au moins en partie le problème du goulet d'étranglement de von Neumann, et avec succès – mais dans certaines limites. Dans la plupart des cas, les améliorations se fondent sur l'idée d'un transfert plus rapide des données de la mémoire vers le bus. Deux principaux facteurs (et plusieurs facteurs secondaires) influent sur la vitesse : la *rapidité de la mémoire* (la vitesse à laquelle la mémoire transfère les données) et la *latence* (le temps nécessaire pour localiser un élément de données particulier).



Comme dans bien d'autres domaines techniques, la médiatisation peut devenir un problème. Le *multithreading*, par exemple, qui consiste à diviser une application ou un ensemble d'instructions en plusieurs unités d'exécution que le processeur traitera une par une, est souvent présenté comme un moyen de résoudre le problème du goulet d'étranglement de von Neumann, alors qu'en réalité il ne fait qu'alourdir la charge de travail (et aggraver le problème). Le multithreading est la réponse à un autre problème : rendre l'application plus efficiente. Quand une application ajoute des problèmes de latence à un goulet d'étranglement de von Neumann, c'est le système tout entier qui est ralenti. Le multithreading permet au processeur d'éviter de perdre du temps à attendre les instructions de l'utilisateur ou de l'application et d'avoir toujours du grain à moudre. Une application peut générer une latence quelle que soit l'architecture du processeur, pas seulement dans le cas d'une architecture de type von Neumann. Dans tous les cas, tout ce qui accélère globalement l'exécution d'une application est visible pour l'utilisateur et le système dans son ensemble.

Le GPU, qu'est-ce que c'est exactement ?

La finalité initiale du processeur graphique (GPU, pour *Graphics Processing Unit*) était le traitement rapide des données d'images suivi de l'affichage à l'écran de l'image résultante. Durant la phase initiale de l'évolution des PC, le processeur central, ou CPU, assurait la totalité du traitement, si bien que l'image pouvait ne s'afficher que lentement pendant que le CPU exécutait d'autres tâches. À ce moment-là, les PC étaient généralement équipés d'un *adaptateur d'affichage*, un dispositif n'ayant aucune ou pratiquement aucune puissance de traitement. Tout ce que fait l'adaptateur d'affichage, c'est convertir les données informatiques sous une forme visuelle. En fait, l'utilisation d'un unique processeur s'est révélée pratiquement infaisable lorsque le PC a cessé de n'afficher que du texte ou des images extrêmement simples en 16 couleurs. Cependant, les GPU n'ont pas représenté une réelle percée dans le domaine de l'informatique avant que l'on commence à vouloir des sorties en 3D. Associer un CPU à un adaptateur d'affichage ne pouvait plus faire l'affaire.

Un premier pas dans cette direction a été réalisé avec des systèmes comme le Hauppauge 4860 (http://www.geekdot.com/hauppauge-4860/), dont la carte mère comportait un CPU et un processeur graphique spécial (le 80860, en l'occurrence). Le 80860 présentait l'avantage d'exécuter les calculs extrêmement vite. Malheureusement, ces systèmes asynchrones multiprocesseurs n'ont pas satisfait aux attentes (même s'ils étaient incroyablement rapides pour l'époque) et ils se sont révélés très onéreux. En outre, se posait le problème de l'écriture d'applications avec prise en compte d'une seconde puce. Les deux puces

partageaient aussi la mémoire (qui était abondante dans ces systèmes).

Avec le GPU, le traitement de l'image est transféré de la carte mère vers la carte graphique périphérique. Le CPU donne au GPU l'instruction d'exécuter une tâche, puis le GPU détermine la meilleure méthode pour l'exécuter indépendamment du CPU. Le GPU dispose d'une mémoire distincte, et le chemin de données pour son bus présente une forte capacité. En outre, le GPU peut accéder à la mémoire principale pour obtenir les données nécessaires à l'exécution d'une tâche et il peut afficher les résultats indépendamment du CPU. C'est donc cette configuration qui permet aujourd'hui l'affichage graphique.



Néanmoins, ce qui distingue réellement le GPU, c'est le fait qu'il comporte généralement plusieurs centaines de cœurs (voir http://www.nvidia.com/object/what-is-gpu-computing.html), tandis qu'un CPU n'en comporte qu'un petit nombre. Même si le CPU offre davantage de fonctionnalités de façon générale, le GPU exécute les calculs de façon extrêmement rapide et peut transférer les données vers l'écran plus rapidement encore. Cette capacité est ce qui fait du GPU un élément essentiel des systèmes actuels.

Pourquoi les processeurs graphiques sont si performants

De même que la puce 80860 présentée dans la section précédente, les processeurs graphiques actuels font merveille dans l'exécution des tâches spécialisées qui sont liées au traitement d'image, notamment le travail avec les vecteurs. Tous ces cœurs qui exécutent des tâches en parallèle accélèrent véritablement les calculs de l'IA.

En 2011, dans le cadre du projet Google Brain (https://research.google.com/teams/brain/), une IA a été programmée pour reconnaître la différence entre les chats et les personnes sur les vidéos de YouTube. Cependant, Google a dû pour cela utiliser 2000 CPU dans un de ses centres de données géants. Rares sont ceux qui disposeraient des ressources nécessaires pour faire la même chose.

Bryan Catanzaro (de l'équipe de recherche de NVIDIA) et Andrew Ng (de Stanford) ont tout de même pu reproduire cette expérience en utilisant une série de 12 processeurs graphiques NVIDIA. Le constat étant fait que les GPU pouvaient remplacer tout un ensemble de systèmes informatiques bourrés de CPU, il a été possible de développer divers projets d'IA. En 2012, en utilisant des processeurs graphiques, Alex Krizhevsky (de l'université de Toronto) a gagné le concours ImageNet de reconnaissance d'image par ordinateur. Un certain nombre de chercheurs ont déjà accompli des merveilles avec des GPU (lire l'article à l'adresse https://adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html).

Créer une configuration de traitement spécialisée

L'apprentissage profond et l'IA sont des processus incompatibles avec une architecture de type von Neumann, aux dires d'un certain nombre de spécialistes comme Massimiliano Versace, P.-D.G. de Neurala Inc. (https://www.neurala.com/). Lorsque la configuration matérielle n'est pas adaptée à la tâche exécutée par l'algorithme, il existe toutes sortes d'inefficiences, il est nécessaire de subdiviser la tâche, et

l'obtention d'un résultat est bien plus difficile qu'elle devrait l'être. Il est donc très souhaitable de concevoir un matériel adapté au logiciel. L'agence de recherche de pointe du ministère de la Défense des États-Unis (DARPA) s'est lancée dans un projet de ce type, avec les *Systems of Neuromorphic Adaptive Plastic Scalable Electronics* (SyNAPSE). L'idée sous-jacente est de reproduire la manière dont la nature résout les problèmes en associant mémoire et puissance de traitement plutôt que de les séparer. Un tel système (d'une taille impressionnante) a effectivement été développé. Pour en savoir plus, consultez l'adresse http://www.artificialbrains.com/darpa-synapse-program.

Le projet SyNAPSE a progressé. IBM a construit un système plus petit en utilisant une technologie moderne à la fois très rapide et économe en énergie électrique (voir http://www.research.ibm.com/cognitive-computing/neurosynaptic-chips.shtml). Le seul problème est que personne ne l'achète. Dans le domaine des cassettes vidéo, Betamax était un système de stockage de données meilleur que VHS, mais c'est VHS qui l'a emporté grâce à un coût moindre, une plus grande facilité d'utilisation et des fonctionnalités attrayantes (voir https://gizmodo.com/betamax-vs-vhshow-sony-lost-the-original-home-video-1591900374). Il s'est passé la même chose avec TrueNorth, l'offre SyNAPSE d'IBM. Il était difficile de trouver une clientèle prête à payer un prix plus élevé, des programmeurs capables de développer un logiciel en utilisant la nouvelle architecture et des produits véritablement plus performants grâce à ce processeur. Ainsi, malgré ses faiblesses inhérentes, la combinaison de CPU et de GPU continue de triompher.

Quelqu'un finira sans doute par mettre au point une puce qui ressemblera davantage à l'équivalent d'un cerveau. Le



système actuel deviendra probablement incapable de permettre les progrès désirés en matière de puissance de traitement. Des compagnies comme Google travaillent actuellement sur des alternatives comme le TPU (*Tensor Processing Unit*, ou unité de traitement de tenseur), déjà utilisé dans des applications comme Google Search, Street View, Google Photos et Google Traduction.

Parce que nous disposons maintenant d'une technologie utilisée pour des applications à grande échelle, certains achètent aussi les processeurs, certains programmeurs savent écrire des applications pour ces processeurs, et il existe des produits attractifs qui sont demandés. Contrairement à SyNAPSE, le TPU repose aussi sur la technologie de l'ASIC (*Application Specific Integrated Circuit*, ou circuit intégré propre à une application), utilisée dans d'innombrables applications, par conséquent Google ne fait en réalité que réaffecter une technologie existante. Il en résulte que ce type de processeur a bien plus de chances de conquérir le marché qu'une chose comme SyNAPSE, qui dépend d'une technologie entièrement nouvelle.

Des matériels de plus en plus performants

Le CPU fait toujours l'affaire pour les systèmes informatiques des entreprises et les applications dans lesquelles la flexibilité dans la programmation, de façon générale, est plus importante que la puissance de traitement en tant que telle. Cependant, les GPU sont maintenant la norme dans différents domaines de la science des données et pour l'apprentissage machine, l'IA et l'apprentissage profond. Naturellement, tout le monde guette toujours la prochaine nouveauté dans l'environnement de développement. Les CPU comme les

GPU sont des processeurs pour le traitement au niveau de la production. Dans l'avenir, il se peut que ces standards soient remplacés par les types de processeurs suivants :

» **Les circuits intégrés propres à une application (ASIC)** : contrairement aux processeurs généralistes, l'ASIC est fabriqué pour une utilisation bien spécifique. Une solution ASIC permet d'obtenir d'excellentes performances avec très peu de consommation électrique, mais la flexibilité est restreinte. Un exemple est le TPU de Google, qui est utilisé pour le traitement de la parole.

» **Les portes programmables par l'utilisateur (FPGA)** : comme pour l'ASIC, un fabricant produit généralement une FPGA pour une application particulière. Cependant, contrairement à l'ASIC, il est possible de programmer la FPGA afin d'en changer la fonctionnalité. Un exemple de solution FPGA est Brainwave, de Microsoft, qui est utilisé dans des projets d'apprentissage https://techcrunch.com/2018/05/07/microsoft-launches-project-brainwave-its-deep-learning-acceleration-platform/.



La bataille entre ASIC et FPGA promet d'être rude, et les développeurs d'IA en seront les vainqueurs. Pour l'instant, Microsoft et FPGA semblent être en tête du classement. La technologie est fluide, et l'on doit s'attendre à de nouveaux développements.

Les fabricants travaillent aussi à mettre au point de nouveaux types de traitements, dont on ne sait s'ils fonctionneront toujours comme prévu. Graphcore, par exemple, travaille sur une unité de traitement intelligent (*Intelligence Processing Unit*, ou IPU).

Concernant ces nouveaux processeurs, il convient de suivre les informations avec un certain recul, compte tenu

du battage dont cette industrie a déjà fait l'objet dans le passé. Quand de grandes compagnies comme Google et Microsoft proposeront réellement des applications d'une nouvelle technologie, on pourra envisager avec un peu plus d'assurance l'avenir de la technologie en question.

Ajouter des capteurs spécialisés

Une caractéristique essentielle de l'IA est sa capacité de simuler l'intelligence humaine à l'aide de toute une série de perceptions. Ce que nous percevons par nos sens est ce qui nous permet de développer les divers types d'intelligence présentés au Chapitre 1. Les sens de l'être humain sont la source adéquate d'intrants à partir desquels on crée un être humain intelligent. Or, même en supposant qu'il devienne possible à une IA de mettre en œuvre ces sept types d'intelligence, il manquera toujours les intrants appropriés pour que l'intelligence soit fonctionnelle.

L'être humain dispose, en principe, de cinq sens pour interagir avec l'environnement : la vue, l'ouïe, le toucher, le goût et l'odorat. Bizarrement, les gens n'ont pas une compréhension totale de leurs capacités, aussi n'est-il pas vraiment surprenant que les ordinateurs soient encore loin de percevoir l'environnement de la même manière que les humains. Ainsi, jusqu'à récemment, le goût ne recouvrait que quatre saveurs : salé, sucré, amer, et acide. Or, deux éléments ont récemment été ajoutés à cette liste : umami et gras (voir l'article en anglais https://fivethirtyeight.com/features/can-we-taste-fat/).

De même, certaines femmes sont tétrachromates (https://concettaantico.com/tetrachromacy/), ce qui signifie qu'elles peuvent distinguer 100000000 couleurs au lieu de 1000000 (seules les femmes peuvent

être tétrachromates, pour une question de chromosomes). Il n'est pas encore possible de savoir combien de femmes ont cette faculté (d'après certaines sources, ce nombre pourrait atteindre 15 %.

Aujourd'hui, l'utilisation de données statiques et dynamiques filtrées permet à une IA d'interagir avec les humains selon des modalités spécifiques. Prenons par exemple Alexa (https://www.amazon.com/Amazon-Echo-And-Alexa-Devices/b?node=9818047011), l'appareil d'Amazon qui vous entend et vous répond. Même si Alexa, en réalité, ne comprend rien de ce que vous dites, la possibilité de communication apparente est très captivante et incite les gens à anthropomorphiser ces appareils. Pour fonctionner, Alexa a besoin d'utiliser un capteur spécial : le microphone, qui lui permet d'entendre. En réalité, Alexa utilise plusieurs microphones afin de capter les sons suffisamment bien pour donner l'illusion d'une compréhension. Malheureusement, aussi avancée qu'elle soit, Alexa ne peut ni voir, ni sentir, ni toucher, ni goûter quoi que ce soit : elle n'est donc humaine en aucune manière.



Dans certains cas, les humains voudraient vraiment que leur IA soit dotée de sens supérieurs ou différents. Une IA peut détecter les mouvements la nuit et réagir en conséquence en utilisant les infrarouges plutôt que la vision normale. Le recours à des sens différents est une des justifications à l'utilisation de l'IA aujourd'hui. La capacité de travailler dans des environnements dans lesquels les humains ne peuvent pas travailler est une des raisons pour lesquelles certains types de robots sont devenus si appréciés, mais le travail dans ces environnements nécessite souvent une série de capteurs non humains. Par conséquent, on peut en réalité distinguer deux catégories de capteurs (qui ne sont ni l'une ni l'autre bien définies) : ceux qui imitent la

perception d'un événement par l'être humain, et ceux qui fonctionnent dans un environnement différent.

Concevoir des méthodes pour interagir avec l'environnement

Une IA qui serait isolée et sans aucune interaction avec l'environnement serait inutile. Bien sûr, cette interaction prend la forme d'intrants et d'extrants. Les méthodes traditionnelles pour fournir des intrants et des extrants consistent à utiliser directement des flux de données compréhensibles par l'ordinateur : jeux de données, requêtes textuelles, *etc*. Or, ces méthodes ne sont pas très conviviales et leur utilisation suppose des compétences particulières.



De plus en plus, les interactions avec une IA prennent une forme plus simple pour les humains que dans le cas d'un contact direct avec l'ordinateur. Quand on pose une question à Alexa, par exemple, l'entrée d'informations (l'intrant) se fait au moyen d'une série de microphones. L'IA transforme les mots-clés contenus dans la question en unités lexicales qu'elle peut comprendre. Ces unités déclenchent alors des calculs qui constituent un extrant. L'IA donne ensuite à l'extrant une forme compréhensible par l'être humain : une phrase parlée. Cette phrase, vous l'entendez quand Alexa s'adresse à vous *via* un haut-parleur. En résumé, si Alexa offre une fonctionnalité pratique, c'est parce qu'elle interagit avec l'environnement de deux manières différentes, qui sont attrayantes pour les humains mais dont elle n'a en réalité aucune compréhension.

Les interactions peuvent prendre différentes formes, et même, un nombre de formes toujours plus grand. Par exemple, une IA peut maintenant sentir les odeurs (voir

http://www.sciencemag.org/news/2017/02/artificial-intelligence-grows-nose). En réalité, l'ordinateur ne sent rien. Ce sont des capteurs qui détectent des substances chimiques et les transforment en données que l'IA utilise ensuite de la même manière que toutes les autres données. La capacité de détecter des substances chimiques n'est pas nouvelle, et la capacité de transformer en chiffres l'analyse de ces substances n'est pas nouvelle non plus, pas plus que les algorithmes utilisés pour interagir avec les données résultantes. Ce qui est nouveau, ce sont les séries de données utilisées pour interpréter les données entrantes comme une odeur, et ces séries de données proviennent d'études effectuées chez les humains. Le nez d'une IA peut avoir toutes sortes d'utilisations. Songeons, par exemple, à la capacité de l'IA à utiliser ce nez dans le contexte de certains environnements dangereux, notamment pour détecter une fuite de gaz avant d'autres types de capteurs.

Les interactions physiques font aussi l'objet de nouveaux développements. Les robots qui travaillent sur les chaînes de montage ne sont plus du tout une nouveauté. Cependant, si maintenant les robots se mettent à conduire les voitures, songeons aux implications que cela peut avoir. L'IA peut aussi agir à une échelle plus modeste. Hugh Herr, par exemple, utilise une IA pour gérer une interaction avec un pied intelligent (voir https://www.smithsonianmag.com/innovation/future-robotic-legs-180953040/). Ce pied dynamique constitue une solution de rechange élaborée pour une personne amputée du pied ou de la jambe. Au lieu de l'impression statique que procure une prothèse classique, il procure une sensation active analogue à celle que nous avons l'habitude d'éprouver en utilisant un vrai pied. La poussée, par exemple, n'est pas la même selon que l'on gravit une côte ou une pente. De même, la

poussée est différente, selon que l'on marche sur un trottoir ou que l'on franchit une marche.

Ce qu'il faut retenir ici, c'est qu'à mesure que l'IA peut exécuter des calculs plus complexes dans des conditionnements plus réduits et avec des ensembles de données toujours plus volumineux, elle est de plus en plus capable d'effectuer des tâches intéressantes. Cependant, à l'heure actuelle, ces tâches ne peuvent pas relever d'une catégorie humaine. Vous ne pourrez pas toujours interagir avec une IA qui comprenne ce que vous dites, mais vous pourrez recourir à une IA pour pouvoir rester en vie ou du moins, pour que votre vie soit plus… vivable.

PARTIE 2
Recenser les utilisations de l'IA dans la société

DANS CETTE PARTIE…

» Travailler avec l'IA dans des applications informatiques

» Utiliser l'IA pour automatiser des processus courants

» Découvrir la manière dont l'IA répond aux besoins de la médecine

» Définir des méthodes pour interagir directement avec les humains

Chapitre 5

Faire le tour des utilisations de l'IA dans les applications informatiques

DANS CE CHAPITRE

» Définir et utiliser l'IA dans des applications

» Utiliser l'IA pour effectuer des corrections et présenter des suggestions

» Comprendre les erreurs potentielles de l'IA

Vous avez sans doute déjà utilisé l'IA sous une forme ou une autre, dans les diverses applications informatiques qui vous servent pour votre travail. Si vous parlez à votre smartphone, par exemple, c'est parce que celui-ci utilise une IA de reconnaissance vocale. C'est aussi une IA qui filtre tout le courrier électronique indésirable qui, autrement, envahirait votre boîte de réception. La première partie de ce chapitre est consacrée aux différents types d'applications de l'IA, dont un certain nombre vous surprendront, et aux domaines dans lesquels on utilise couramment l'IA pour effectuer un grand nombre de tâches. Vous y découvrirez aussi une source de limitations à la création d'applications de l'IA, ce qui vous permettra de comprendre pourquoi il n'existera probablement jamais de robots doués de conscience – du

moins, jamais avec les technologies dont nous disposons actuellement.

Cependant, que l'IA puisse ou non parvenir un jour à la capacité d'éprouver quelque chose, toujours est-il qu'elle peut accomplir un grand nombre de tâches utiles. Pour l'essentiel, l'IA répond actuellement à nos besoins de deux manières : les corrections et les suggestions. Il ne s'agit pas d'interpréter ces deux termes littéralement. Une correction n'est pas nécessairement la réponse à une erreur, et une suggestion n'est pas nécessairement la réponse à une requête. Prenons l'exemple d'une voiture à conduite assistée (lorsque l'IA assiste le conducteur plutôt que de le remplacer). Au cours du trajet, les corrections apportées par l'IA seront fonction des conditions de circulation et de l'état des routes, de la présence de piétons et de tout un ensemble d'autres problèmes éventuels à résoudre à l'avance. L'IA adopte une approche proactive d'un problème qui ne se produira pas nécessairement. De même, l'IA peut suggérer au conducteur une certaine trajectoire présentant la plus grande chance d'être pertinente, et la suggestion pourra changer par la suite en fonction des nouvelles conditions de circulation. La deuxième partie de ce chapitre traite des corrections et des suggestions de façon séparée.

La troisième partie de ce chapitre est consacrée aux erreurs potentielles de l'IA. Une erreur se produit chaque fois que le résultat est différent de ce qui était attendu. Il se peut qu'un résultat positif soit inattendu. Naturellement, des erreurs pures et simples peuvent aussi se produire : il peut arriver qu'une IA ne produise pas un résultat positif, ou même, que le résultat obtenu soit en contradiction avec l'objectif initial (avec des dégâts comme conséquence possible). Il ne faut pas perdre de vue que les résultats des applications de l'IA ne sont pas nécessairement tout blancs ou tout noirs, mais qu'il convient d'envisager des nuances de gris, si l'on veut comprendre l'impact de l'IA

sur des applications informatiques qui produisent un résultat absolument correct ou absolument incorrect.

Avoir une idée des applications les plus courantes

La diversité des applications possibles de l'IA ne connaît comme limites que celles de l'imagination du programmeur, et la plupart des possibilités n'ont pas encore été découvertes. En effet, compte tenu de la flexibilité qu'offre l'IA, ses applications peuvent voir le jour ailleurs que dans le cadre dans lequel le programmeur les a définies initialement. Il se peut qu'un de ces jours, un logiciel d'IA programme lui-même sa prochaine génération (voir https://www.technologyreview.com/s/603381/ai-software-learns-to-make-aisoftware/). Cependant, pour avoir une meilleure idée de l'utilité de l'IA dans les applications informatiques, il est utile de passer en revue ses utilisations actuellement les plus courantes (ainsi que les pièges éventuels qui leur sont associés), et c'est l'objet des sections qui suivent.

Utiliser l'IA dans des applications types

Vous pourrez rencontrer l'IA là où il était difficile de l'imaginer. Si le thermostat intelligent qui régule la température de votre habitation est un modèle très élaboré, il se peut qu'il incorpore une IA (voir https://www.popsci.com/gadgets/article/2011-12/artificially-intelligent-thermostats-learns-adaptautomatically). Le recours à l'IA, même dans ces applications très particulières, est justifié

lorsqu'il s'agit de ce que l'IA sait faire le mieux, comme par exemple enregistrer les températures qui sont préférées dans le temps afin de créer automatiquement un programme de températures. Voici des applications de l'IA parmi les plus typiques et les plus répandues :

- la créativité artificielle ;
- la vision par ordinateur, la réalité virtuelle, et le traitement d'image ;
- les diagnostics ;
- la reconnaissance des visages ;
- les jeux avec IA, les jeux sur ordinateur, la théorie des jeux, et la planification stratégique ;
- la reconnaissance d'écriture manuscrite ;
- le traitement automatique du langage naturel, la traduction automatique, et les agents conversationnels ;
- le contrôle non linéaire et la robotique ;
- la reconnaissance optique de caractères ;
- la reconnaissance de la parole.

Prendre la mesure du vaste champ d'application de l'IA

L'IA est utilisée de différentes manières selon les applications. L'IA est aussi utilisée de façon plus générique dans certains domaines d'expertise. La liste suivante recense les domaines dans lesquels l'IA intervient le plus souvent :

- la vie artificielle ;
- le raisonnement automatisé ;

» l’automatisation ;
» l’informatique d’inspiration biologique ;
» l’extraction de concepts ;
» l’exploration de données ;
» les filtres contre le courrier indésirable ;
» les systèmes intelligents hybrides ;
» l’agent intelligent et le contrôle intelligent ;
» la représentation des connaissances ;
» la résolution des contentieux ;
» la robotique : robotique comportementale, cognition, cybernétique, robotique développementale (épigénétique) et robotique évolutionnaire ;
» le Web sémantique.

Analyser l’argument de la chambre chinoise

En 1980, un article rédigé par John Searle et intitulé « Minds, Brains, and Programs » (Esprits, cerveaux et programmes) avait été publié dans la revue *Behavioral and Brain Sciences*. Il s’agissait d’une réfutation du test de Turing, dans lequel un ordinateur pouvait tromper un être humain en imitant le comportement d’une personne, au moyen d’une série de questions (lire l’article, traduit en français, à l’adresse https://www.persee.fr/doc/quad_0987-1381_1987_num_1_1_2097). L’hypothèse de base était que le fonctionnalisme, ou la capacité de simuler des

caractéristiques spécifiques de l'esprit humain, diffère de la pensée véritable.

L'argument de la chambre chinoise, puisque c'est le nom donné à cette expérience de pensée, repose sur deux tests. Dans le premier test, on crée une IA capable d'interpréter une question sous forme de caractères chinois, d'utiliser une série de règles pour préparer une réponse, puis de produire cette réponse sous forme de caractères chinois. La question concerne une histoire : l'IA doit interpréter les questions qui lui sont soumises de telle sorte que la réponse ne soit pas aléatoire mais fasse référence à l'histoire concernée. L'IA est assez performante pour que personne, en dehors de la chambre, ne puisse savoir que c'est une IA qui accomplit les tâches requises. Une personne qui parle le chinois croira vraiment que l'IA est capable de lire et de comprendre cette langue.

Dans le second test, une personne humaine qui ne connaît pas le chinois reçoit trois éléments qui représentent ce que fait l'ordinateur. Le premier est une table de consultation constituée d'un grand nombre de caractères chinois, le deuxième est une histoire en chinois, et le troisième est une série de règles servant à établir les liens entre le premier élément et le deuxième. On soumet à la personne une batterie de questions écrites en chinois et elle se réfère aux règles pour trouver dans l'histoire une réponse, d'après une interprétation des caractères. La réponse est la série de caractères chinois qui correspond à la question d'après les règles. La personne en question réussit si bien cette tâche, que personne ne peut la soupçonner de ne pas comprendre le chinois.

L'objet de ces deux tests est de démontrer que la capacité d'utiliser des règles formelles pour produire un résultat (une syntaxe) n'a rien à voir avec la compréhension de ce qui est dit ou exécuté (la sémantique). Searle postulait que

la syntaxe ne suffit pas pour qu'il existe une sémantique, or c'est ce que les promoteurs de l'IA affirment lorsqu'il s'agit de fabriquer des systèmes fonctionnant à partir de règles comme le Script Applier Mechanism (SAM) : voir https://eric.ed.gov/? id=ED161024.

La question sous-jacente est de savoir s'il s'agit d'une IA forte, qui comprend réellement ce qu'elle vise à réaliser, ou d'une IA faible, qui ne fait que suivre les règles. Aujourd'hui, il n'existe que des IA faibles : aucune IA n'est en mesure de comprendre quoi que ce soit. Il s'agit d'une simulation réalisée au moyen d'une programmation subtile, et qui applique des règles (comme celles qui sont implicites dans les algorithmes). Naturellement, l'idée que quelle que soit la complexité future des machines, elles ne développeront pas un cerveau, ce qui revient à dire qu'elles n'acquerront jamais une compréhension des choses, suscite une importante controverse. D'après Searle, l'IA restera faible. Ce sujet fait l'objet d'une discussion sur la page http://www.iep.utm.edu/chineser/. Les arguments et contre-arguments sont intéressants à lire, car ils fournissent des explications très utiles concernant l'enjeu réel de la création d'une IA.

Découvrir comment l'IA rend une application plus conviviale

La question de la convivialité d'une application grâce à l'IA peut s'envisager sous différents angles. Au niveau le plus élémentaire, l'IA peut anticiper l'entrée de l'utilisateur. Par exemple, quand celui-ci vient juste de taper les premiers caractères d'un mot, l'IA devine les caractères restants. De ce fait, l'IA répond à plusieurs objectifs :

» Parce qu’il n’a plus besoin de saisir que quelques caractères, l’utilisateur réalise un gain de temps et d’efficacité.

» L’application reçoit moins d’entrées erronées, car il y a moins de fautes de frappe.

» L’utilisateur et l’application communiquent mieux, sachant que l’application présente à l’utilisateur des termes corrects ou améliorés dont il pouvait ne pas se souvenir, ce qui évite la saisie de termes différents que la machine risquerait de ne pas reconnaître.

Une IA peut aussi apprendre à partir des saisies précédemment effectuées par l’utilisateur, en réorganisant les suggestions d’une manière adaptée à la méthode de travail de l’utilisateur. Le niveau suivant d’interaction relève du domaine des suggestions, un sujet abordé dans la section « Formuler des suggestions », plus loin dans ce chapitre. Les suggestions peuvent aussi consister à fournir à l’utilisateur des idées qui ne lui seraient peut-être pas venues à l’esprit autrement.



Même dans le domaine des suggestions, on pourrait avoir tendance à croire que l’IA pense, bien que ce ne soit pas le cas. L’IA exécute une forme avancée de reconnaissance des formes et d’analyse pour déterminer la probabilité du besoin d’un intrant particulier. La section « Analyser l’argument de la chambre chinoise », précédemment dans ce chapitre, étudie la différence entre l’IA faible, le type d’IA que l’on trouve dans les applications actuelles, et l’IA forte, que certaines applications pourront peut-être intégrer un jour dans l’avenir.

L’utilisation d’une IA signifie aussi qu’il est désormais possible de mettre en œuvre d’autres sortes d’entrées intelligentes. L’exemple de la voix est presque galvaudé, mais c’est une des méthodes les plus courantes.

Cependant, même si l'IA ne présente pas toute la gamme de sens évoqués au Chapitre 4, elle peut fournir une grande variété d'entrées intelligentes non verbales. Un choix évident est la perception visuelle, comme la reconnaissance du visage du propriétaire ou d'une menace en fonction de l'expression faciale. L'entrée peut toutefois inclure un système de contrôle des signes vitaux de l'utilisateur, par exemple, en vue de détecter d'éventuels problèmes. En fait, une IA pourrait utiliser un nombre considérable d'entrées intelligentes qui, pour la plupart, restent encore à inventer.

Actuellement, les applications ne prennent généralement en compte que ces trois premiers niveaux de convivialité. Cependant, à mesure des progrès de l'IA, il devient essentiel pour une IA de présenter des comportements d'intelligence artificielle amicale (IAA) cohérents avec une intelligence artificielle générale (IAG) ayant un impact positif sur l'humanité. L'IA a des objectifs, mais qui ne sont pas nécessairement en phase avec une éthique humaine, et le risque de discordance est aujourd'hui source d'inquiétude. Une IAA doit inclure une logique qui permette que ses objectifs restent cohérents avec ceux de l'humanité, de manière similaire aux trois lois que l'on trouve dans les ouvrages d'Isaac Asimov (https://www.auburn.edu/~vestmon/robotics.html) et qui sont abordées plus en détail au Chapitre 12. Cependant, il est souvent considéré que ces trois lois ne constituent qu'un bon point de départ (http://theconversation.com/after-75-yearsisaac-asimovs-three-laws-of-robotics-need-updating-74501) et qu'il nous faut d'autres garde-fous.



Naturellement, toute cette discussion sur les lois et l'éthique peut se révéler assez déroutante et difficile à cerner. Un exemple simple d'IAA serait le refus de divulguer des informations personnelles sur l'utilisateur,

sauf celles que le destinataire a besoin de connaître. En fait, une IAA pourrait même aller plus loin en appliquant la reconnaissance de formes aux intrants d'origine humaine, en y localisant les informations susceptibles de présenter un caractère personnel, et en avertissant l'utilisateur du potentiel de dommages avant tout envoi de l'information vers une destination donnée. En résumé, une IA peut changer significativement la manière dont les gens envisagent les applications et interagissent avec elles.

Effectuer automatiquement des corrections

Les humains ne cessent de tout corriger. Ce n'est pas parce que tout va mal, mais plutôt parce que l'on recherche toujours des améliorations possibles. Même lorsque l'on a réussi à atteindre le niveau de pertinence ou d'excellence voulu à un moment donné, une nouvelle expérience amène à remettre ce résultat en question, sachant que l'on dispose maintenant de données supplémentaires à partir desquelles on juge ce résultat dans une situation particulière. Pour imiter vraiment l'intelligence humaine, l'IA doit aussi avoir cette capacité de corriger constamment les résultats qu'elle produit, même si ces résultats apportent déjà une réponse positive. Les sections qui suivent traitent du problème de savoir ce qui est correct ou adéquat, et étudient la manière dont les corrections automatisées peuvent parfois échouer.

Distinguer les différentes sortes de corrections

La plupart des gens, si on leur parle de corrections effectuées par une IA, penseront au vérificateur d'orthographe ou au correcteur grammatical. Disposer d'une IA pour détecter et corriger les fréquentes erreurs d'origine humaine dans les documents est certes une bonne idée.

Les corrections peuvent prendre toutes sortes de formes, et il ne s'agit pas nécessairement d'erreurs commises ou pouvant être commises dans l'avenir. Dans une voiture, par exemple, l'IA peut assister le conducteur en corrigeant constamment la trajectoire. Même si l'automobiliste a une conduite sûre, l'IA peut procéder à des microcorrections pour que le trajet soit plus sûr encore.

Pour aller plus loin, supposons que la voiture qui arrive en face s'arrête soudainement parce qu'un animal se trouve sur la chaussée. Même si notre conducteur n'a commis aucune erreur, l'IA peut réagir plus vite que lui et arrêter la voiture aussi vite que possible, et de la façon la plus sûre.

Apprécier les avantages des corrections automatiques

Quand une IA détecte un besoin de correction, elle peut demander à l'utilisateur humain la permission d'effectuer la correction, ou bien, elle peut procéder au changement de façon automatique. Par exemple, quand l'utilisateur se sert d'un outil de reconnaissance de la parole pour écrire un document et fait une faute de grammaire, l'IA doit lui demander la permission de corriger, sachant qu'elle n'a pas nécessairement compris ce qu'il voulait.

En revanche, il est parfois très important que l'IA accomplisse un processus de décision assez sûr pour pouvoir effectuer des corrections automatiquement. Si l'on reprend l'exemple de la section précédente, dans lequel il

fallait que la voiture freine, l'IA n'a pas le temps de demander la permission au conducteur : il faut freiner immédiatement, sans quoi la collision pourrait être mortelle. Dans le domaine de l'IA, les corrections automatiques ont lieu d'être, à condition qu'une décision s'impose d'urgence et que l'IA soit assez fiable et efficace.

Comprendre pourquoi les corrections automatiques ne font pas l'affaire

En réalité, comme on l'a vu dans la section « Analyser l'argument de la chambre chinoise », précédemment dans ce chapitre, une IA ne comprend rien du tout. Or, sans compréhension, il n'y a pas de capacité de faire face à des circonstances imprévues. En l'occurrence, il s'agit d'un événement impromptu, à propos duquel l'IA ne peut pas accumuler des données supplémentaires ni recourir à une autre solution d'ordre mécanique. Un être humain pourra résoudre le problème, sachant qu'il comprend en quoi ce problème consiste et qu'il a généralement une perception du contexte assez satisfaisante pour pouvoir définir un schéma aboutissant à une solution. Par ailleurs, l'innovation et la créativité dont l'être humain peut faire preuve sont sources de solutions qui ne seraient pas évidentes par d'autres moyens. Sachant que pour l'instant l'IA ne peut pas faire preuve d'innovation ni de créativité, elle n'est pas bien placée pour résoudre des problèmes dans des domaines spécifiques.

Pour mettre les choses en perspective, prenons le cas d'un vérificateur d'orthographe. Il se peut que l'utilisateur saisisse un mot tout à fait correct, même s'il ne figure pas dans le dictionnaire utilisé par l'IA pour les corrections. Souvent, l'IA remplace le mot saisi par un mot qui lui

ressemble, mais qui reste incorrect. Même une fois que l'utilisateur a vérifié le document, a retapé le mot correct, puis l'a ajouté au dictionnaire, l'IA risque encore de faire une erreur. Elle pourra, par exemple, traiter l'abréviation *CPU* différemment de *cpu* parce que la première est en majuscules et la seconde en minuscules. Une personne humaine saurait que ces deux abréviations n'en sont qu'une et que dans le second cas, la saisie est correcte même si elle devrait peut-être s'afficher en majuscules.

Formuler des suggestions

Une suggestion n'est pas la même chose qu'une commande. Même si certaines personnes ne semblent pas le comprendre, une suggestion est simplement une idée proposée comme solution éventuelle à un problème. Formuler une suggestion signifie que d'autres solutions pourraient exister, et accepter une suggestion ne signifie pas automatiquement qu'elle sera appliquée. En fait, la suggestion n'est qu'une idée qui peut même ne pas être pertinente. Bien sûr, dans un monde idéal, toutes les suggestions seraient pertinentes : elles seraient au moins des solutions possibles pour un résultat correct, ce qui est rarement le cas dans le monde réel. Les sections qui suivent décrivent la nature des suggestions applicables à une IA.

Obtenir des suggestions basées sur des actions passées

Pour une IA, le moyen le plus courant de formuler une suggestion consiste à recenser des actions passées, en tant qu'événements, puis à se fonder sur ces actions en tant que données. Supposons qu'une personne achète un gadget chaque début de mois, pendant trois mois. Il sera

alors logique de lui suggérer d’en acheter un autre au début du quatrième mois. Une IA vraiment performante devrait formuler la suggestion à la bonne date. Par exemple, si l’utilisateur a toujours effectué son achat entre le 3 et le 5 du mois, l’IA devrait commencer à formuler la suggestion le troisième jour du mois, et passer à autre chose au-delà du cinquième jour.

Un utilisateur humain produit un grand nombre d’indices en accomplissant des tâches. Contrairement aux humains, une IA prend note de chacun de ces indices et les enregistre de façon cohérente. C’est ainsi qu’elle peut ensuite, dans de nombreux cas, formuler des suggestions en fonction des actions passées avec un degré élevé de pertinence.

Obtenir des suggestions basées sur les groupes

Une autre technique courante pour formuler des suggestions consiste à se référer à l’appartenance à un groupe. Cette appartenance n’a pas besoin d’être formelle. Un groupe est une association informelle de personnes ayant en commun quelque besoin ou activité d’importance mineure. Parmi les amateurs de romans policiers, par exemple, on peut imaginer trouver aussi bien un bûcheron qu’un commerçant ou un diététicien. Même s’ils n’ont rien d’autre en commun, pas même leur lieu de résidence, par le fait même qu’ils aiment tous les trois les romans policiers, ils font partie d’un groupe. Une IA peut facilement détecter ce genre de propriété commune qu’une personne humaine n’aurait peut-être jamais remarquée, ce qui lui permet de formuler des suggestions d’achat appropriées.

Un groupe peut être constitué de liens immatériels et fugaces, temporaires dans le meilleur des cas. Ainsi, par exemple, toutes les personnes ayant emprunté le vol 1982 au départ de Houston à une certaine date pourraient constituer un groupe. Là encore, il n'existe entre ces personnes aucun lien d'aucune sorte en dehors du fait que leurs noms apparaissent sur la liste des passagers d'un vol donné. Cependant, grâce à la connaissance de cette information, une IA pourrait procéder à un filtrage additionnel afin d'identifier, parmi les passagers, ceux qui aiment les romans policiers. L'idée est qu'une IA peut produire de bonnes suggestions sur la base de l'appartenance à un groupe, même si ce groupe est difficile (voire impossible) à identifier du point de vue humain.

Obtenir des suggestions non pertinentes

Quiconque a déjà effectué des achats en ligne sait que les sites Internet affichent souvent des suggestions en fonction de différents critères, notamment les achats antérieurs. Malheureusement, ces suggestions sont souvent non pertinentes car l'IA sous-jacente n'est pas douée de compréhension. Le fait que l'achat d'un certain article soit exceptionnel dans la mesure où l'on ne peut, en principe, avoir besoin de cet article qu'une fois dans sa vie, est compréhensible pour un humain, mais pas pour une IA. Par conséquent, à moins que le programmeur n'ait prévu une règle particulière spécifiant que ce type d'article ne s'achète qu'une fois, l'IA risque de continuer à recommander le même produit, sachant que le volume des ventes est limité. En suivant une règle secondaire consistant à promouvoir les produits qui se vendent le moins bien, l'IA se comporte selon les caractéristiques

prévues par le développeur, mais les suggestions qu'elle émet ne sont pas pertinentes du tout.

En dehors des erreurs de l'IA liées aux règles et des erreurs logiques, les suggestions peuvent aussi devenir non pertinentes en raison d'un problème de données. Un GPS, par exemple, peut proposer un itinéraire en fonction des meilleures données possibles pour une certaine destination, mais il est possible aussi que l'itinéraire proposé soit impraticable parce qu'une route est barrée pour cause de travaux. Certes, les applications de GPS prennent souvent en compte les travaux d'entretien des routes, mais d'autres problèmes peuvent ne pas être pris en considération, comme un abaissement soudain de la limite de vitesse, ou des conditions météorologiques qui rendent une route dangereuse. L'être humain peut pallier l'insuffisance de données grâce à une innovation, par exemple en optant pour l'itinéraire le moins fréquenté ou en sachant interpréter les panneaux de déviation.

Même quand une IA n'est pas mise en défaut par des problèmes de logique, de règles ou de données, elle peut formuler de mauvaises suggestions faute de comprendre la corrélation entre certaines séries de données, de la même manière qu'une personne humaine. Ainsi, par exemple, si vous achetez de la tuyauterie et des carreaux de plâtre pour effectuer des travaux de plomberie, il est évident pour vous qu'à l'issue de ces travaux il sera nécessaire de repeindre la zone concernée. L'être humain est doué d'un sens esthétique qui fait défaut à l'IA, et il établit une corrélation entre différents produits qui n'est pas à la portée de l'IA.

Étudier les erreurs de l'IA

On parlera d'erreur pure et simple quand le résultat d'un processus, compte tenu des inputs, n'est correct en aucun

cas, c'est-à-dire quand la réponse n'est pas adaptée du tout à la requête. Les exemples d'erreurs d'IA sont difficiles à trouver. Un article récent de *BBC News* expliquait comment une différence d'un seul pixel sur une image pouvait induire en erreur une IA particulière (lire l'article à l'adresse http://www.bbc.com/news/technology-41845878). À propos de l'impact des attaques contradictoires contre l'IA. L'article de *Kaspersky Lab Daily*, à l'adresse https://www.kaspersky.com/blog/ai-fails/18318/, présente d'autres exemples de situations dans lesquelles une IA ne donne pas une réponse correcte. L'idée est que l'IA présente toujours un taux d'erreurs élevé dans certaines circonstances, et que les développeurs qui travaillent avec l'IA ne savent généralement pas très bien pourquoi des erreurs se produisent.

Dans le domaine de l'IA, les sources d'erreurs sont nombreuses. Cependant, comme on l'a vu au Chapitre 1, l'IA n'étant pas capable d'émuler les sept formes d'intelligence humaine, les erreurs sont non seulement possibles, mais inévitables. Le Chapitre 2 est consacré en grande partie aux données et à l'impact sur l'IA d'une imperfection dans les données. Le Chapitre 3 vous révèle aussi que même les algorithmes utilisés par l'IA présentent des limites. Le Chapitre 4 explique qu'une IA n'imite pas tous les sens dont l'être humain est doté.

Un problème majeur qui devient de plus en plus visible est qu'avec l'IA, les entreprises ont tendance à escamoter les problèmes, ou même, à les ignorer. On recourt à l'IA surtout pour réduire les coûts et accroître la productivité, et ce n'est pas toujours une solution viable. L'article de Bloomberg à l'adresse https://www.bloomberg.com/news/articles/2017-06-13/the-limits-of-artificial-

intelligence traite de ce problème en détail. Un des exemples récents les plus intéressants d'entreprise ayant abusé de l'IA est celui de Tay, une IA lancée par Microsoft (lire l'article à l'adresse https://www.geekwire.com/2016/microsoft-chatbot-tay-mit-technology-fails/) qui a appris à diffuser des propos racistes, sexistes et obscènes auprès d'un vaste public au cours d'une présentation.



Ce qu'il convient avant tout de retenir de cette section, ce n'est pas que l'IA ne serait pas fiable ou serait inutilisable. En réalité, une IA peut permettre à un humain compétent de travailler plus rapidement et plus efficacement. Elle peut permettre de limiter le nombre d'erreurs courantes ou répétitives. Dans certains cas, les erreurs de l'IA peuvent même égayer la journée en apportant un peu d'humour. Néanmoins, l'IA ne pense pas, et dans un certain nombre de situations dynamiques, elle ne peut pas encore remplacer l'être humain. L'IA est intéressante surtout quand une personne humaine contrôle ses décisions, ou quand l'environnement est assez statique pour que de bons résultats aient une probabilité élevée d'être obtenus (du moins, à condition que personne ne s'emploie à induire l'IA en erreur).

Chapitre 6

Automatiser des processus courants

DANS CE CHAPITRE

» **Utiliser l'IA pour servir les besoins des humains**

» **Rendre l'industrie plus efficiente**

» **Mettre au point des protocoles de sécurité dynamiques à l'aide de l'IA**

Le Chapitre 5 étudie l'utilisation de l'IA dans une application, une situation dans laquelle une personne humaine interagit avec l'IA de façon pertinente et concrète, même si cette personne ne sait pas qu'elle a affaire à une IA. La finalité est de permettre à la personne d'accomplir quelque chose plus rapidement, plus facilement, plus efficacement, ou de répondre aux besoins d'un tiers. Un *processus* qui inclut une IA est différent d'un processus qui n'en inclut pas, car l'IA assiste un utilisateur humain ou accomplit une autre tâche sans intervention humaine directe. La première section de ce chapitre traite de la manière dont ces processus aident les humains. Sachant que l'ennui est peut-être la pire des situations d'un point de vue humain (qu'il suffise de penser à tout ce qui peut se produire de négatif quand les gens s'ennuient), ce chapitre aborde le processus de l'IA sous l'angle de l'ennui.

Quand l'IA est utilisée dans un processus, c'est souvent dans l'industrie. Il suffit de penser à tous les robots qui équipent aujourd'hui les usines dans le monde entier. Même lorsque l'automatisation à base d'IA remplace l'être humain, elle est aussi un facteur de sécurité car elle exécute des tâches qui sont généralement considérées comme dangereuses. Assez curieusement, l'ennui est une des principales causes des accidents dans l'industrie et la source d'un certain nombre d'autres problèmes (voir https://thepsychologist.bps.org.uk/volume-20/edition-2/boredom-work). Or, les robots peuvent exécuter toutes ces tâches répétitives pendant un temps prolongé, et sans jamais s'en lasser.

Au cas où vous ne vous seriez pas encore suffisamment ennuyé, vous pouvez aussi lire quelque chose sur ce sujet dans la troisième section de ce chapitre, à propos des nouveaux domaines dans lesquels l'IA excelle et de la manière dont l'environnement de travail, dans tous les cas, devient plus sécurisé. En fait, rien que dans l'industrie automobile, les progrès qu'apporte l'utilisation de l'IA sont multiples et variés, entre autres dans le domaine de la sécurité (voir https://igniteoutsourcing.com/publications/artificial-intelligence-in-automotive-industry/).



L'idée essentielle de ce chapitre est que l'IA est utile dans les processus, surtout là où les humains ont tendance à s'ennuyer, avec pour conséquence qu'ils font des erreurs qu'une IA a peu de chances de commettre. Naturellement, une IA ne peut pas éliminer toutes les sources de perte d'efficience, de désintérêt et de risque. L'utilisateur peut choisir d'ignorer l'aide de l'IA, mais les limites sont bien plus profondes encore. Comme on l'a vu dans les chapitres précédents (et plus particulièrement au Chapitre 5), une IA n'est pas douée de compréhension : elle ne peut pas trouver des solutions créatives ou

innovantes, par conséquent certains problèmes ne peuvent pas être résolus grâce à l'IA, quels que soient les efforts consacrés à son élaboration.

Élaborer des solutions contre l'ennui

Les sondages montrent souvent davantage ce que les gens croient vouloir que ce qu'ils veulent vraiment, et cependant ils sont utiles. De jeunes diplômés ont été interrogés sur le style de vie auquel ils aspiraient, et aucun n'a parlé de l'ennui (voir https://www.huffingtonpost.com/paul-raushenbush/what-kind-of-life-do-you_b_595594.html). En fait, en sondant pratiquement n'importe quel groupe de personnes, on a peu de chances d'obtenir une réponse dans laquelle il est question d'ennui. La plupart des gens n'ont pas envie de s'ennuyer (nous ne dirons pas que c'est le cas de tout le monde, ou nous risquerions de recevoir une avalanche de courriers électroniques avec des contre-exemples). Dans certains cas, l'IA peut travailler avec les humains pour rendre la vie plus intéressante, du moins, pour l'être humain. Les sections qui suivent traitent des solutions que peut apporter l'IA contre l'ennui (et de quelques-unes qu'elle ne peut pas apporter).

Rendre les tâches plus intéressantes

Toute occupation, qu'elle concerne un individu ou une organisation, présente certaines caractéristiques qui font qu'elle attire des gens et leur donne envie de participer. À l'évidence, certaines occupations, comme prendre soin de

ses propres enfants, même si elles ne sont pas rémunérées, peuvent être extrêmement gratifiantes. Inversement, le métier de comptable, par exemple, peut être bien payé et cependant, ne pas apporter une grande satisfaction professionnelle. Divers sondages (par exemple sur la page http://www.careercast.com/jobs-rated/jobs-rated-report-2016-ranking-200-jobs) et articles (comme sur la page http://www.nytimes.com/2010/09/12/jobs/12search.html) portent sur l'arbitrage entre rémunération financière et satisfaction, mais leur lecture peut souvent induire en erreur, car le fondement sur lequel on peut prendre une décision est ambigu. Néanmoins, il ressort de toutes ces sources qu'à partir d'un certain niveau de rémunération, l'épanouissement personnel devient essentiel dans l'intérêt suscité par l'activité en question (quelle qu'elle soit). Bien sûr, il est pratiquement impossible de définir précisément ce qui apporte de la satisfaction, mais le fait que l'activité soit intéressante reste un critère important. Une activité intéressante présentera toujours un potentiel de satisfaction élevé.



Le problème n'est pas nécessairement de changer de travail, mais de rendre le travail plus intéressant pour éviter l'ennui. Une IA peut effectivement faciliter ce changement, en éliminant les répétitions. Cependant, des exemples comme Alexa, d'Amazon, et Home, de Google, montrent qu'il existe des alternatives. Le sentiment de solitude qui peut régner à la maison, sur le lieu de travail, en voiture ou ailleurs est une importante source d'ennui. Quand on se sent seul, la dépression guette, et l'ennui n'est souvent que l'étape préliminaire. Créer des applications utilisant l'interface d'Alexa (voir https://developer.amazon.com/) ou Actions sur Google API (voir https://developers.google.com/actions/) pour

simuler des interactions humaines appropriées permet de rendre le lieu de travail plus attrayant. Plus important, développer des interfaces intelligentes de ce type peut permettre aux humains d’accomplir toutes sortes de tâches triviales rapidement, comme rechercher des informations ou interagir avec des appareils intelligents au lieu de simplement appuyer sur des boutons (voir https://store.google.com/product/google_home).

Permettre aux gens de travailler plus efficacement

La plupart des gens, ou du moins la plupart de ceux qui sont tournés vers l’avenir, ont des idées sur la manière dont ils aimeraient voir une IA leur faciliter la vie en éliminant les tâches auxquelles ils n’ont pas envie de se consacrer. Il s’agit souvent de tâches banales, mais certaines méritent de retenir notre attention, comme détecter une insatisfaction chez son conjoint et envoyer des fleurs. Ce n’est sans doute pas réalisable, mais l’idée n’en est pas moins intéressante.

Le fait est que les gens proposeront sans doute les idées les plus intéressantes sur la façon de créer une IA qui réponde spécifiquement aux besoins de la personne. Dans la plupart des cas, les idées sérieuses seront appropriées également pour d’autres utilisateurs. L’automatisation des dossiers d’incident, par exemple, est une chose qui pourrait être applicable à un certain nombre de secteurs d’activité. Si quelqu’un pouvait mettre au point une interface générique avec une partie programmable pour produire des dossiers d’incident personnalisés adéquats, l’IA ferait gagner un temps précieux aux utilisateurs et permettrait des gains d’efficience futurs, car les dossiers enregistreraient toujours l’information requise.

Comprendre comment l'IA permet d'éviter l'ennui

L'ennui apparaît sous de nombreuses formes, et ce sentiment est perçu de différentes manières. Il y a l'ennui qui découle de l'absence des ressources et des connaissances nécessaires, ou du fait que d'autres besoins ne soient pas satisfaits. Un autre type d'ennui provient du fait de ne pas savoir quoi faire. Une IA peut changer les choses dans le premier cas, mais pas dans le second. Cette section est consacrée au premier cas (et la section suivante au second).



L'accès à des ressources de toutes sortes permet de réduire l'ennui en permettant aux humains de créer sans l'ennuyeuse nécessité d'acquérir le matériel nécessaire. Voici ce que peut notamment faire une IA pour faciliter l'accès aux ressources :

- chercher en ligne les articles nécessaires ;
- commander automatiquement les articles nécessaires ;
- assurer le contrôle des capteurs et autres systèmes d'acquisition de données ;
- gérer les données ;
- exécuter des tâches triviales ou répétitives.

Comprendre comment l'IA ne permet pas d'éviter l'ennui

Comme on l'a vu dans les chapitres précédents, et plus particulièrement dans les Chapitres 4 et 5, une IA n'est ni créative ni intuitive. Par conséquent, si vous demandez à

une IA de penser à quelque chose que vous pourriez faire, il y a peu de chances que cela donne des résultats satisfaisants. On pourrait programmer l'IA pour qu'elle identifie les dix choses que vous aimez le plus faire et qu'elle en choisisse ensuite une au hasard, mais le résultat ne sera toujours pas satisfaisant car l'IA ne pourra pas prendre en compte des aspects comme votre état d'esprit sur le moment. En réalité, même avec l'expression faciale la plus convaincante, une IA ne sera pas capable d'interagir avec vous d'une manière susceptible de produire un résultat satisfaisant.

Une IA ne peut pas non plus vous motiver. Songez à ce qui se produit quand un ami cherche à vous motiver (ou quand vous cherchez à motiver cet ami). L'ami s'appuie sur une combinaison de connaissances intrapersonnelles (empathie à l'idée de ce qu'il ressentirait dans votre situation) et interpersonnelles (en projetant des idées créatives sur le moyen d'obtenir de vous une réaction émotionnelle positive). Une IA ne peut pas acquérir le premier type de connaissance et ne peut acquérir le second type de connaissance qu'en quantités extrêmement limitées, comme on l'a vu au Chapitre 1. Par conséquent, une IA ne peut pas réduire votre ennui par des techniques motivationnelles.



Quoi qu'il en soit, l'ennui n'est pas nécessairement une mauvaise chose. Un certain nombre d'études récentes montrent que l'ennui favorise la pensée créative (voir par exemple https://www.fastcompany.com/3042046/the-science-behind-how-boredom-benefitscreative-thought), et c'est bien dans cette direction que les humains doivent progresser. Après avoir passé en revue d'innombrables articles qui donnent à croire que l'IA va faire disparaître des emplois, il est important de réaliser que les emplois en question sont souvent ennuyeux et ne laissent pas aux personnes qui

les occupent du temps pour créer. Aujourd'hui encore, les gens pourraient trouver des emplois productifs et créatifs s'ils y songeaient vraiment. L'article « 7 Surprising Facts About Creativity, According To Science » (https://www.fastcompany.com/3063626/7-surprising-facts-about-creativity-according-to-science) traite du rôle des activités ennuyeuses et de la rêverie dans la stimulation de la créativité. Dans l'avenir, si les humains veulent réellement atteindre les étoiles et faire d'autres choses extraordinaires, la créativité sera essentielle. Que l'IA ne puisse rien contre votre ennui, c'est donc une bonne chose.

Travailler dans un contexte industriel

Dans un contexte industriel, il existe toujours des situations dangereuses, quels que soient le temps, les efforts et l'argent que l'on consacre à ce problème. Vous trouverez facilement des articles comme celui-ci, qui présente sept types de dangers courants en milieu industriel : http://www.safetyandhealthmagazine.com/articles/14054-commonworkplace-safety-hazards. Ces problèmes ont bien souvent une cause humaine et l'ennui ne fait que les aggraver, mais le fait est que l'environnement dans lequel les gens travaillent pose réellement beaucoup de problèmes. Les sections suivantes expliquent comment l'automatisation peut permettre aux gens de vivre mieux et plus longtemps.

Développer différents niveaux d'automatisation

L'automatisation dans l'industrie est bien plus ancienne qu'on pourrait le penser. On croit souvent qu'elle a commencé avec la chaîne de montage de Henry Ford (voir http://www.history.com/this-day-in-history/fords-assembly-lines-tarts-rolling). En réalité, l'automatisation a commencé en 1104 à Venise (voir https://www.mouser.com/applications/factory-automation-trends/), où 16000 ouvriers ont été capables de construire entièrement un bateau de guerre en une seule journée. Les Américains ont répété cet exploit durant la Seconde Guerre mondiale (voir https://www.nps.gov/nr/travel/wwiibayarea/shipbuilding.htm). L'automatisation, c'est donc une longue histoire.

Ce qui est bien plus récent, c'est une IA réellement capable d'assister les humains dans le processus d'automatisation. Aujourd'hui, bien souvent, c'est d'abord l'opérateur humain qui définit la manière d'exécuter la tâche, définit le travail à faire, puis le délègue à l'ordinateur. Un exemple est celui de l'automatisation des processus robotiques (PAR), qui permet d'apprendre à un logiciel à se comporter comme un humain dans l'utilisation d'applications. Ce processus diffère de la programmation telle qu'elle s'effectue en utilisant par exemple Visual Basic pour les applications (VBA) dans Office, dans la mesure où l'application est spécifique et ne nécessite pas de codage. Beaucoup de gens sont surpris d'apprendre qu'il existe dix degrés d'automatisation, dont neuf peuvent dépendre d'une IA. Le degré que vous choisissez dépend de votre application :

1. **Un opérateur humain crée une tâche et la confie à un ordinateur.**
2. **Une IA aide l’humain à déterminer les alternatives.**
3. **L’IA détermine les meilleures alternatives, puis l’humain doit accepter ou refuser la recommandation.**
4. **L’IA détermine les alternatives, les utilise pour définir une série d’actions, puis soumet cette liste d’actions à un humain qui doit accepter ou refuser chaque action avant la mise en œuvre.**
5. **L’IA détermine les alternatives, définit une série d’actions, crée une tâche, puis demande à l’humain de la valider avant de la soumettre à l’ordinateur.**
6. **L’IA crée automatiquement la tâche et l’envoie à la file d’attente des tâches de l’ordinateur, l’opérateur humain jouant le rôle d’intermédiaire dans le cas où la tâche sélectionnée devrait être terminée avant la mise en œuvre effective.**
7. **L’IA crée et met en œuvre la tâche, puis informe l’opérateur humain de ce qui a été fait si la tâche nécessite une correction ou une reprise.**
8. **L’IA crée et met en œuvre la tâche, puis informe l’opérateur humain de ce qui a été fait, mais seulement si l’humain le lui demande.**
9. **L’IA crée et met en œuvre la tâche sans en rendre compte, sauf dans le cas où l’humain a besoin d’intervenir, par exemple si une erreur est survenue ou si le résultat n’est pas celui attendu.**
10. **L’IA initie le besoin de la tâche au lieu d’attendre que ce soit l’humain qui lui demande de la créer.**

Elle n’en rend compte que dans le cas où l’humain a besoin d’intervenir, par exemple si une erreur est survenue. L’IA peut produire un degré donné de correction d’erreur et gère elle-même les résultats inattendus.

Utiliser autre chose que seulement des robots

Quand il s’agit de l’industrie, la plupart des gens pensent à l’automatisation, aux tâches confiées à des robots. Or, notre société en est à sa quatrième révolution industrielle : nous avons eu la machine à vapeur, puis la production de masse, puis l’automatisation, et maintenant c’est la communication (voir https://www.engineering.com/ElectronicsDesign/ElectronicsDesignArticles/ArticleID/8379/New-Chips-are-Bringing-Factory-Automation-into-the-Era-of-Industry-40.aspx). Une IA a besoin d’informations provenant de toutes sortes de sources pour pouvoir effectuer des tâches de façon efficiente. Il s’ensuit que plus les informations que l’on peut obtenir sont nombreuses et variées, mieux l’IA peut travailler (en supposant que les données sont gérées correctement). Sachant cela, toutes sortes d’industries recourent aujourd’hui à un moteur de communication industrielle (ICE) pour coordonner la communication entre les diverses sources que nécessite une IA.

Dans l’industrie, les robots accomplissent une grande partie du travail, mais il faut aussi des capteurs pour évaluer les risques potentiels, par exemple liés aux tempêtes. Cependant, la coordination devient plus importante encore pour que l’activité reste efficiente. Faire en sorte que les camions livrant les matières premières arrivent en temps et en heure et que d’autres camions

soient disponibles au moment voulu pour enlever les produits finis, par exemple, ce sont là des tâches essentielles au bon fonctionnement des entrepôts. L'IA doit connaître le programme d'entretien de tous les équipements, afin de faire en sorte que chaque équipement soit entretenu de la meilleure manière possible (afin d'accroître la fiabilité) et que les opérations d'entretien s'effectuent au moment où cet équipement est le moins sollicité (afin d'accroître l'efficience). L'IA doit aussi prendre en compte des problèmes comme le coût des ressources. Ainsi, par exemple, il est parfois avantageux de faire fonctionner certains équipements le soir, quand l'électricité est moins chère.

Compter seulement sur l'automatisation

Dans les premiers exemples d'usines automatisées, un environnement spécial était nécessaire : dans les fabriques de puces électroniques, par exemple, les locaux devaient être particulièrement propres. Depuis, l'automatisation s'est généralisée. En raison des dangers pour les humains et du coût de l'utilisation de la main-d'œuvre pour certaines tâches industrielles, on trouve aujourd'hui de nombreux exemples d'installations dans lesquelles aucune intervention humaine n'est nécessaire.



À un moment donné, un certain nombre de technologies rendront possible l'exécution de toutes les tâches d'un processus industriel sans intervention humaine (pour des exemples, voir https://www.automationmag.com/opinion/thought-leaders/5248-top-10-industrialautomation-trends-automationdirect). L'idée est qu'au bout du compte, il faudra que la société trouve des tâches à confier aux

humains, qui seront des tâches différentes des tâches répétitives que comportent les processus industriels.

Créer un environnement sécurisé

Un des rôles les plus souvent évoqués de l'IA, en dehors de l'automatisation des tâches, est celui d'assurer la sécurité des gens de diverses manières. Des articles comme celui de la page https://futurism.com/7-reasons-you-should-embrace-not-fear-artificial-intelligence/ décrivent un environnement dans lequel l'IA joue un rôle d'intermédiaire et encaisse les coups à la place des humains quand survient un problème de sécurité. La sécurité revêt toutes sortes de formes. Certes, l'IA permettra de réduire les risques au travail dans des environnements variés, mais elle permettra aussi de créer un environnement plus sûr et de réduire les risques associés à des tâches courantes, comme surfer sur l'Internet. Les sections qui suivent proposent une vue générale de la façon dont une IA peut permettre de disposer d'un environnement plus sûr.

Prendre en compte l'ennui comme cause d'accidents

Au volant comme au travail, l'ennui accroît la fréquence de toutes sortes d'accidents. En fait, quand une personne qui exécute une tâche nécessitant un certain niveau d'attention a tendance à somnoler, les conséquences sont rarement heureuses. Ce problème est si grave et si important qu'il existe pléthore d'articles sur ce sujet. En voici un exemple : « Modelling human boredom at work : mathematical formulations and a probabilistic framework » (http://www.emeraldinsight.com/doi/full/10.

1108/1741038131132798l). La survenue d'un accident (ou le fait qu'un accident ne soit évité que de justesse) est un événement aléatoire. On pourrait imaginer que l'on mette au point des algorithmes permettant de déterminer la probabilité d'accidents provoqués par l'ennui dans certaines situations.

Envisager l'IA pour éviter les problèmes de sécurité

Aucune IA ne peut empêcher que surviennent des accidents dont les causes sont humaines, comme l'ennui. Dans le meilleur des cas, lorsque les humains s'appliquent à respecter les règles que l'IA permet d'établir, l'IA ne peut que permettre d'éviter des problèmes potentiels. En matière de sécurité, aucun environnement ne bénéficie des Trois lois de la robotique définies par Asimov : c'est à nous autres humains de décider de ne pas nous mettre en danger. Voici ce que peut faire l'IA dans ce domaine :

- Suggérer des rotations de postes (sur le lieu de travail, au volant, ou même à la maison) pour que les tâches restent stimulantes.
- Contrôler l'activité humaine afin de suggérer des temps morts de façon plus pertinente, pour cause de fatigue ou autre.
- Assister les utilisateurs dans l'exécution des tâches pour associer l'intelligence d'origine humaine à la rapidité de réaction de l'IA.
- Accroître les capacités de détection de la présence humaine pour que les problèmes de sécurité éventuels soient plus évidents.
- Prendre en charge les tâches répétitives pour que les humains risquent moins d'être fatigués et

> participent surtout aux aspects les plus intéressants du travail.

Se rendre compte que l'IA ne peut pas faire disparaître les problèmes de sécurité

Garantir une sécurité totale suppose la faculté de voir l'avenir. Comme l'avenir est inconnu, les risques potentiels pour les personnes à un moment donné sont inconnus également, sachant qu'une situation imprévue peut toujours se produire. Une situation imprévue est une situation que les auteurs d'une stratégie de sécurité particulière n'ont pas envisagée. Si nous sommes doués pour trouver de nouveaux moyens de nous retrouver dans des situations difficiles, nous autres humains, c'est en partie parce que nous sommes curieux et créatifs. Il est dans la nature humaine de trouver une méthode pour contourner les mesures de sécurité assurées par une IA : nous voulons voir ce qui va se produire si nous essayons quelque chose, généralement quelque chose de stupide.

Les situations imprévues ne sont pas le seul problème auquel une IA se trouve confrontée. Même si l'on trouvait tous les moyens possibles de se mettre en danger, la puissance de traitement requise pour détecter l'événement et déterminer une ligne de conduite à adopter serait astronomique. L'IA travaillerait si lentement que sa réponse interviendrait toujours trop tard pour être d'une quelconque utilité. En conséquence, quand un équipement de sécurité doit faire intervenir une IA pour assurer le niveau de sécurité requis, ses concepteurs doivent se fonder sur des calculs de probabilités et envisager une protection contre les événements qui risquent le plus de se produire.

Chapitre 7

Utiliser l’IA pour répondre à des besoins médicaux

DANS CE CHAPITRE

» **Mieux suivre les patients**

» **Assister les humains dans des tâches variées**

» **Analyser les besoins des patients**

» **Réaliser des opérations chirurgicales et d’autres tâches spécialisées**

La médecine, c’est compliqué. Si la formation d’un médecin peut durer jusqu’à une quinzaine d’années, voire davantage, selon la spécialité, ce n’est pas sans raison (voir http://work.chron.com/long-become-doctor-us-7921.html). Quand un médecin termine ses études en ayant acquis suffisamment de connaissances pour pouvoir espérer faire carrière, la plupart des gens sont déjà dans le monde du travail depuis onze ans (la majorité des gens ayant arrêté leurs études au bac ou à l’issue d’un cycle court). En parallèle, l’apparition de nouvelles technologies et de nouvelles méthodologies contribue à rendre les tâches toujours plus complexes. À un moment donné, il devient impossible d’être compétent, même dans une spécialité bien circonscrite. Naturellement, c’est là une des principales raisons pour lesquelles l’irremplaçable être humain a besoin de l’aide cohérente, logique et objective que peut

lui apporter une IA. Il s'agit d'abord d'apporter au médecin, dans le suivi de ses patients (c'est le sujet de la première section de ce chapitre), une aide qu'il serait impossible à envisager de la part d'une personne humaine compte tenu du nombre élevé de contrôles, de l'impérieuse nécessité de les effectuer dans un certain ordre et d'une manière particulière, et du potentiel d'erreurs qui est monumental.

Heureusement, les patients ont aujourd'hui davantage de possibilités que jamais auparavant d'accomplir eux-mêmes un certain nombre de tâches relevant des processus thérapeutiques. Il en est ainsi du recours à des jeux, par exemple, sachant que le patient est guidé par une application qui vérifie qu'il accomplit la tâche de manière appropriée pour pouvoir guérir dans les meilleures conditions. Les progrès réalisés dans le domaine des prothèses et autres aides médicales permettent aussi aux patients d'être moins dépendants de l'assistance des professionnels. La deuxième section de ce chapitre explique comment une IA peut permettre au patient de satisfaire ses propres besoins en matière de soins médicaux.

De même qu'il s'avère difficile, sinon impossible, de corriger des dispositifs sans les voir fonctionner dans un environnement spécifique, de même les humains s'opposent parfois à l'analyse nécessaire pour diagnostiquer les problèmes. Des possibilités d'analyse variées permettent au médecin d'identifier un trouble particulier et de le traiter plus facilement. Aujourd'hui, il est tout à fait possible pour le médecin d'équiper un patient d'un dispositif de surveillance, d'effectuer un suivi à distance et de recourir à une IA pour effectuer l'analyse permettant de diagnostiquer un problème, tout cela sans que le patient ait besoin d'effectuer plus d'une visite (pour la mise en place du dispositif de surveillance). Dans certains cas, par exemple l'utilisation d'un glucomètre, le patient peut même aller acheter l'appareil en magasin, si

bien que la visite chez le médecin n'est même plus nécessaire.

Même si la troisième section de ce chapitre ne dit rien des systèmes d'analyse, elle vous permet d'avoir une bonne vue d'ensemble du sujet. Bien sûr, il se peut que le patient ait impérativement besoin d'une intervention chirurgicale ou d'une autre procédure (voir la quatrième section de ce chapitre), mais une solution robotique peut parfois exécuter la tâche mieux que le médecin. Dans certains cas, une assistance robotisée permet au médecin d'être plus efficace et de concentrer son attention là où seul l'être humain peut agir. Le recours à divers types de technologies rend aussi le diagnostic plus facile, plus rapide et plus précis. Ainsi, par exemple, une IA peut permettre au médecin de localiser le début d'un cancer bien plus tôt.

Mettre en œuvre une surveillance portative pour le patient

Un professionnel de la santé n'est pas toujours capable de diagnostiquer un problème de santé chez un patient simplement en écoutant son cœur, en contrôlant ses signes vitaux ou en faisant pratiquer une analyse de sang. Le corps n'émet pas toujours des signaux donnant des informations utiles. En outre, certaines fonctions corporelles, comme le taux glycémique, changent d'un moment à un autre, si bien qu'un contrôle constant devient nécessaire. Or, si vous alliez consulter le médecin à chaque fois qu'un de ces contrôles est nécessaire, ce serait coûteux en temps et pas forcément très utile. Pour déterminer certaines caractéristiques corporelles, les anciennes méthodes exigeaient une intervention manuelle externe de la part du patient, un processus qui était source d'erreurs. C'est pour ces raisons, et pour bien d'autres

encore, qu'une IA peut permettre de suivre les données d'un patient d'une manière efficiente, moins propice aux erreurs et plus cohérente, comme l'expliquent les sections qui suivent.

Porter des moniteurs utiles

Il existe toutes sortes de moniteurs utiles. En réalité, un certain nombre n'ont rien à voir avec la profession médicale, mais produisent des résultats positifs pour la santé du patient. C'est le cas, par exemple, du moniteur Moov (https://welcome.moov.cc/), qui contrôle à la fois votre rythme cardiaque et votre mouvement en 3D. L'IA associée à cet appareil relève ces données et vous fournit des conseils pour améliorer votre entraînement, par exemple en vous indiquant la manière dont vos pieds entrent en contact avec le trottoir quand vous courez, ou en vous disant si vous devez allonger le pas. Ce genre d'appareil vous permet de vous assurer que votre entraînement est bénéfique et que vous ne risquez pas de vous blesser.

Au cas où un appareil de contrôle de type bracelet serait encore trop volumineux, Motiv (https://mymotiv.com/) propose un anneau de plus petite dimension, qui mesure à peu près autant de données que Moov. Il surveille même votre sommeil et vous aide à mieux vous reposer la nuit. Ce genre d'anneau présente généralement une série d'avantages et d'inconvénients. Pour en savoir plus, lisez l'article à l'adresse https://www.wareable.com/fashion/best-smart-rings-1340. Comme vous le verrez sur les photos de ce site Internet, ces accessoires ressemblent davantage à des bijoux qu'à des appareils de contrôle : vous avez à la fois la santé et l'élégance dans un même objet.

Bien sûr, si votre seul objectif est de contrôler votre rythme cardiaque, votre choix pourra se porter sur Apple Watch (https://www.apple.com/fr/watch/), qui offre les mêmes possibilités d'analyse à l'aide d'une IA. Tous ces appareils communiquent avec votre smartphone. Vous pouvez donc relier vos données à d'autres applications, ou bien les envoyer à votre médecin.

Utiliser des moniteurs portables essentiels

Un problème dans le domaine de la santé est que l'état du patient peut changer à tout moment, si bien qu'un contrôle intermittent ne fait généralement pas l'affaire. La glycémie, une statistique mesurée par les diabétiques, en est un bon exemple. En surveillant souvent dans la journée les fluctuations du taux de glucose, il devient plus facile d'adapter son traitement médicamenteux et son style de vie, de manière à maîtriser son diabète. Des appareils comme K'Watch (https://www.pkvitality.com/fr/ktrack-glucose/) permettent un contrôle constant, en association avec une application mobile permettant de disposer d'informations utiles pour gérer le diabète. Bien sûr, cela fait des années que sont utilisés des appareils de contrôle intermittent, mais celui-ci offre un degré supplémentaire de contrôle qui permet de réduire nettement les nuisances qu'entraîne cette pathologie.

Surveiller constamment sa glycémie ou une donnée relative à une autre maladie chronique peut sembler exagéré, mais il y a tout de même une utilité pratique à cela. Des produits comme Sentrian (https://sentrian.com.au/) permettent d'utiliser les données distantes pour prédire et devancer l'apparition d'une maladie chez un patient. Grâce à des changements

apportés à la médication et au comportement du patient avant que le mal ne survienne, Sentrian permet de réduire le nombre d’hospitalisations inévitables : ainsi, les patients vivent bien mieux, et les coûts des traitements médicaux diminuent.



LES APPAREILS MÉDICAUX ET LA SÉCURITÉ

Un problème que posent les technologies médicales de toutes sortes est le manque de sécurité. L'idée que quelqu'un puisse pirater un appareil implanté dans votre organisme est effrayante. L'article de la page http://www.independent.co.uk/life-style/gadgets-and-tech/news/first-onlinemurder-will-happen-by-end-of-year-warns-us-firm-9774955.html décrit ce qui pourrait se produire dans un tel cas. Heureusement, d'après un certain nombre de sources, personne n'est encore mort à cause de cela.

Imaginez cependant que votre pompe à insuline ou votre défibrillateur implanté se dérègle par suite d'un piratage, et songez aux dégâts que cela pourrait entraîner. Aux États-Unis, la Federal Drug Administration (FDA) a publié des directives sur la sécurité des appareils médicaux, comme l'explique l'article de la page http://www.securityweek.com/fda-releasesguidance-medi-cal-device-cybersecurity, mais il semble que ces directives ne soient pas appliquées. Cet article explique aussi que les fabricants essaient de trouver des moyens de leurs appareils.

L'IA n'est pas responsable du manque de sécurité que présentent ces appareils, mais en cas d'accident, il se pourrait bien qu'elle soit montrée du doigt. Quand on utilise une IA, il s'agit d'en considérer tous les aspects, surtout quand la santé humaine est en jeu comme dans le cas d'un appareil médical implantable.

Certains équipements sont vraiment essentiels, comme le défibrillateur portable WDV (*Wearable Defibrillator Vest*), qui contrôle continuellement l'état cardiaque et délivre un choc si jamais le cœur cesse de fonctionner correctement (voir https://www.healthafter50.com/heart-health/article/wearable-defibrillator-vest-pros-and-cons). Cette solution à court terme peut permettre à un médecin de décider si le patient a besoin d'une version intracorporelle de ce dispositif. Il présente des avantages et des inconvénients, mais quand il s'agit de sauver une vie, la possibilité de bénéficier sur-le-champ d'un choc salutaire n'a pas de prix. Ce que cet appareil a de plus précieux, c'est le contrôle qu'il permet d'avoir. Tous les patients n'auront pas besoin d'un dispositif implantable, et un tel système de contrôle est essentiel pour éviter une intervention chirurgicale non nécessaire.

Utiliser des moniteurs mobiles

Aujourd'hui, sur le marché, le nombre et la variété des appareils de surveillance médicale fonctionnant grâce à l'IA sont impressionnants. Vous pouvez, par exemple, acheter une brosse à dents électrique intégrant une IA qui surveillera vos habitudes de brossage et qui vous conseillera une meilleure technique de brossage (http://www.businesswire.com/news/home/20170103005546/en/Kolibree-Introduces-Ara-Toothbrush-Artificial-Intelligence/). Quand on y pense, la conception d'un tel appareil pose un certain nombre de difficultés, dont la moindre n'est pas d'arriver à faire en sorte que la partie contenant les circuits électroniques puisse être introduite dans la bouche. Bien sûr, certains pourraient penser que le brossage des dents et la bonne santé sont deux choses différentes, mais ils auraient tort (voir

https://www.mayoclinic.org/healthy-lifestyle/adult-health/in-depth/dental/art-20047475).

Quand un moniteur est mobile, cela signifie généralement qu'il est à la fois miniaturisé et peu intrusif. La simplicité est aussi une exigence pour les appareils conçus pour être utilisés par des personnes qui n'ont pas de compétence en matière médicale. Un exemple est l'électrocardiogramme portable (ECG). Chez le médecin, l'appareil du patient est relié par câbles à un appareil semi-portable qui effectue le monitoring nécessaire. Le QardioCore permet l'ECG sans utilisation de câbles, et une personne ayant peu de connaissances en médecine peut facilement l'utiliser. Cet appareil, comme beaucoup d'autres, communique avec votre smartphone pour fournir l'analyse voulue et assurer des connexions vers des sources extérieures le cas échéant.



Les appareils médicaux actuels sont performants, mais pas portables. L'intérêt de disposer d'une application dotée d'IA et d'appareils spécialisés est d'obtenir des données essentielles au moment où l'on en a besoin, plutôt que de devoir les attendre. Même si vous n'achetez pas une brosse à dents électrique pour contrôler votre technique de brossage ni un ECG pour surveiller votre cœur, sachant que ces appareils sont petits, performants et faciles à utiliser, vous en profiterez peut-être à un moment ou à un autre.

Rendre les gens plus capables

Les techniques actuelles pour permettre aux gens de rester plus longtemps en bonne santé, plutôt que d'accroître simplement l'espérance de vie, sont généralement liées à la possibilité de les rendre davantage capables d'améliorer leur propre santé de

différentes manières. Vous trouverez de nombreux articles qui vous indiqueront 30, 40 voire 50 moyens d'y parvenir, mais bien souvent il s'agit surtout de manger sainement, de faire suffisamment d'exercice physique de manière appropriée et de bien dormir. Bien sûr, il est pratiquement impossible de déterminer le régime alimentaire, le type d'exercice physique et la technique de sommeil qui vous conviendront le mieux. Les sections suivantes traitent des moyens par lesquels un dispositif à base d'IA pourrait vous permettre de bien vivre pendant 80 ans ou plus, plutôt que 60 ans seulement (en réalité, on trouve facilement aujourd'hui des articles qui promettent une espérance de vie de mille ans ou plus dans l'avenir, grâce au progrès des technologies).

Une thérapie par les jeux

Une console de jeu peut être un outil de thérapie physique performant et distrayant. Les consoles Nintendo Wii et Xbox 360 sont utilisées dans un certain nombre de programmes de physiothérapie (https://www.webpt.com/blog/post/do-you-wii-hab-using-motion-gamingyour-therapy-clinic). L'objectif est de faire faire à la personne des mouvements particuliers. Comme dans le cas d'un joueur normal, le jeu récompense automatiquement les bons mouvements du joueur, mais la thérapie a l'avantage d'être distrayante. De ce fait, le patient a plus de chances de vraiment jouer le jeu, c'est le cas de le dire, et de guérir plus vite.

Certes, le mouvement à lui seul, même en utilisant le jeu qui convient, n'est pas une garantie de succès. En jouant à ces jeux, on peut aussi se faire une nouvelle blessure. Avec l'extension Jintronix de Kinect pour Xbox, l'utilisation de cette console de jeu à des fins thérapeutiques se généralise

(https://www.forbes.com/sites/kevinanderton/2017/09/30/jintronix-program-uses-xbox-kinect-hardware-to-help-rehab-patients-infographic/), ce qui augmente la probabilité d'un excellent résultat.

BIAIS, SYMPATHIE ET EMPATHIE

Être bien soigné est l'objectif initial de quiconque pénètre dans un centre de soins. Le traitement dont le patient va bénéficier est censé être le meilleur, mais il est aussi censé être juste. Ce que l'IA peut apporter dans le domaine médical, c'est l'assurance que les compétences techniques restent fortes et qu'il n'y ait aucune partialité d'aucune sorte, du moins, pas du point de vue de l'IA.

Les humains présenteront toujours un biais, car l'être humain est doué d'intelligence intrapersonnelle (voir Chapitre 1). Même la personne la mieux intentionnée et la plus altruiste présentera un certain biais, généralement inconscient, si bien que le prestataire de soins et le patient ne verront pas la même chose (voir la section « Tenir compte des cinq types de données incorrectes » au Chapitre 2). Cependant, le patient en aura presque certainement conscience, et compte tenu de son état, l'affront involontaire sera amplifié. C'est là un problème que le recours à une IA pour assurer l'équité dans la prise en charge du patient permet d'éviter. L'IA peut aussi permettre aux dispensateurs de soins de découvrir des inexactitudes (délibérées ou non) de la part des patients dans leur manière d'indiquer leurs symptômes, et ainsi, d'améliorer les soins.

Ce qui est parfois problématique en médecine, c'est le fait que les compétences techniques ne suffisent pas. Les gens se plaignent souvent du manque d'égards de la part du personnel médical. Ces mêmes patients qui veulent un traitement équitable veulent aussi que le personnel soignant fasse preuve d'empathie (avec pour conséquence que les soins ne seront plus équitables car il y aura alors un biais). L'empathie diffère de la sympathie en contexte. Une personne fait preuve d'*empathie* quand elle se montre capable de ressentir les choses de la même manière que le patient et de créer un cadre de référence qui soit commun avec lui. Deux exercices dans la section « Envisager des solutions logicielles » de ce chapitre

vous permettent de comprendre comment il est possible de créer ce cadre de référence générateur d'empathie. Une IA n'en sera jamais capable, car elle ne pourra pas posséder la nécessaire conscience d'un sentiment, ni comprendre la nécessité de créer un cadre de référence, et elle ne possédera pas l'intelligence intrapersonnelle qui permet d'utiliser ce cadre de référence.

Malheureusement, l'empathie peut aveugler le dispensateur de soins et lui faire perdre de vue les véritables besoins en matière de soins, s'il fausse ainsi sa perspective en ne voyant plus les choses que du point de vue du patient. C'est pourquoi le personnel soignant préfère souvent faire preuve de *sympathie*, la sympathie consistant à porter un regard extérieur et à comprendre ce que le patient peut ressentir (plutôt que ce que le patient ressent effectivement), sans créer un cadre de référence. Ainsi, le personnel soignant peut offrir au patient le soutien émotionnel dont celui-ci a besoin, mais en même temps rester conscient de la nécessité d'effectuer des actes que le patient risque de ne pas apprécier sur le court terme. Une IA ne peut pas accomplir cette tâche, faute de posséder l'intelligence intrapersonnelle et de comprendre assez bien le concept de perspective pour pouvoir l'appliquer de manière appropriée.

Envisager l'utilisation d'exosquelettes

Pour une IA, une des applications les plus complexes consiste à fournir un soutien au corps humain tout entier. C'est ce qui se produit quand une personne porte un exosquelette (autant dire, un robot portable). Une IA détecte les mouvements (ou la nécessité d'un mouvement) et produit une réponse adaptée. Les militaires ont excellé dans l'utilisation d'exosquelettes (voir http://exoskeletonreport.com/2016/07/milita

ry-exoskeletons/). Imaginez que vous puissiez courir plus vite et porter des charges bien plus lourdes grâce au port d'un exosquelette. La vidéo à l'adresse https://www.youtube.com/watch?v=p2W23ysgWKI vous donne un petit aperçu de ce qui est possible. Naturellement, les militaires continuent à expérimenter ces équipements, qui trouvent des applications civiles. L'exosquelette que vous verrez apparaître un jour sur le marché (vous pouvez être pratiquement sûr que cela se produira à un moment donné) sera sans doute un équipement d'origine militaire.

L'industrie a aussi adopté la technologie des exosquelettes (voir par exemple https://www.nbcnews.com/mach/science/new-exoskeleton-does-heavy-lifting-factoryworkers-ncna819291). Les ouvriers des usines sont actuellement exposés à diverses maladies à cause des blessures liées aux gestes répétitifs. Par ailleurs, le travail en usine est extrêmement fatigant. Le port d'un exosquelette diminue la fatigue, mais réduit aussi la fréquence des erreurs et rend les ouvriers plus efficaces. Une personne qui conserve ses niveaux d'énergie au cours de la journée peut en faire plus tout en risquant moins de se blesser, d'endommager des produits ou de blesser quelqu'un d'autre.

Les exosquelettes qui sont utilisés dans l'industrie aujourd'hui reflètent leurs débuts militaires. Les capacités et l'apparence de ces appareils vont évoluer, et ils ressembleront davantage aux exosquelettes que l'on peut voir dans des films comme *Aliens* (https://www.amazon.com/exec/obidos/ASIN/B01I0K018W/datacservip0f-20/). Les exemples réels de cette technologie (voir par exemple la vidéo et l'article sur la page http://www.bbc.com/news/technology-

26418358) sont un peu moins impressionnants, mais ils continueront à gagner en fonctionnalités.

Aussi intéressante que puisse être l'utilisation des exosquelettes pour conférer d'étonnantes capacités aux personnes valides, ce qu'ils peuvent permettre de faire aux personnes à mobilité réduite est plus incroyable encore. Ainsi, un article publié récemment par l'institution Smithsonian traite de l'utilisation d'un exosquelette pour permettre à un enfant souffrant de paralysie cérébrale de marcher (https://www.smithsonianmag.com/innovation/this-robotic-exos-keleton-helps-kids-cerebral-palsywalk-upright-180964750/). Cependant, les exosquelettes utilisés dans des applications médicales ne sont pas toujours utilisables à vie. Par exemple, un exosquelette peut permettre à une personne victime d'un AVC de marcher normalement à nouveau, mais une fois que cette personne a recouvré en grande partie ses facultés, l'exosquelette devient moins utile. À la fin, elle n'en a plus besoin. Certains utilisateurs ont même combiné leur exosquelette avec d'autres produits comme Alexa, d'Amazon.



De manière générale, on ne porte pas un exosquelette pour devenir Iron Man. Il s'agit plutôt de limiter les blessures dues aux gestes répétitifs et de permettre aux gens d'exceller dans des tâches qui, actuellement, se révèlent trop épuisantes ou trop éprouvantes pour l'organisme. D'un point de vue médical, l'utilisation d'un exosquelette est avantageuse parce qu'elle permet aux patients de conserver plus longtemps leur mobilité, or la mobilité est indispensable pour rester en bonne santé.

IMAGINER LE CÔTÉ OBSCUR DES EXOSQUELETTES

Des recherches approfondies sur Internet n'ont pas permis de trouver des utilisations répréhensibles des exosquelettes, à moins d'en condamner les applications militaires. Il est pourtant plus facile de détruire que de créer. Un jour ou l'autre, quelqu'un imaginera une utilisation malveillante d'un exosquelette (ou de n'importe laquelle des technologies évoquées dans ce chapitre). Imaginons, par exemple, des gangsters de haut vol qui utiliseraient avec profit des exosquelettes pour voler des objets lourds.

Si ce livre se propose de dissiper les malentendus qui concernent l'IA et d'en présenter des utilisations positives, il est cependant difficile de s'intéresser à des technologies sans au moins réfléchir un minimum à leur éventuel « côté sombre ». Ce n'est certes pas sans danger lorsque l'on tire la sonnette d'alarme sans qu'aucun fait tangible ne le justifie. Oui, des gens malhonnêtes pourraient faire un usage détourné des exosquelettes, et cela devrait inciter leurs promoteurs à les sécuriser, mais rien de ce genre ne s'est encore produit. Le développement d'une technologie comme l'IA s'accompagne toujours de considérations éthiques relatives à ses utilisations potentielles, positives et négatives.

Ce livre présente diverses considérations éthiques et morales concernant une utilisation positive de l'IA, pour le bien de la société. Sans aucun doute, il est important de sécuriser la technologie, mais il ne faut pas oublier qu'éviter une technologie en raison de son potentiel négatif est véritablement contre-productif.

Répondre à des besoins particuliers

Il fut un temps où un handicap particulier comme la perte d'un membre impliquait des consultations médicales répétées sur plusieurs années, des capacités réduites, et une vie plus brève et moins heureuse. Dans de nombreux cas, il n'en est plus ainsi aujourd'hui, grâce aux progrès réalisés dans le domaine des prothèses et autres dispositifs, lesquels fonctionnent souvent avec le concours de l'IA.



Les besoins particuliers de certaines personnes sont souvent assimilés à une déficience physique ou mentale, ou même à un handicap. Or, chacun de nous a des besoins particuliers. À la fin d'une longue journée, il peut être profitable pour une personne dont la vision est parfaitement normale d'utiliser un zoom logiciel pour voir le texte ou l'image en plus gros. Un outil de conversion de couleurs peut permettre à un utilisateur qui voit normalement les couleurs de remarquer des détails qui, normalement, ne sont pas visibles (du moins, pour lui). Quand on avance en âge, on a généralement davantage besoin d'une assistance pour entendre, pour voir, pour toucher, pour interagir avec des éléments courants de son environnement. De même, une aide à la marche, par exemple, peut permettre à quelqu'un de finir sa vie chez lui au lieu d'être obligé de déménager dans une maison de soins. En résumé, des technologies à base d'IA peuvent permettre aux gens d'avoir une meilleure vie, comme le montrent les sections qui suivent.

ÉVITER DE DIRE N'IMPORTE QUOI SUR LES BESOINS PARTICULIERS

Ces gens dont nous entendons parler de temps à autre, qui semblent particulièrement doués pour mener une vie palpitante malgré leurs besoins particuliers, sont généralement des personnes hors du commun. Ils ont travaillé vraiment dur pour en arriver là. L'utilisation d'un appareil équipé d'une IA peut vous ouvrir des portes, mais pour aller plus loin, il faut que vous ayez la volonté de faire tout ce qu'il faut pour obtenir un résultat. En général, il s'agit de suivre une thérapie en faisant preuve de persévérance. Ce chapitre n'est pas destiné à faire bon marché de la quantité admirable d'efforts que ces personnes remarquables ont réalisés pour vivre mieux. Il s'agit plutôt de mettre l'accent sur les technologies qui contribuent à rendre ces accomplissements possibles. Si vous voulez vraiment voir quelque chose d'admirable, jetez un coup d'œil à l'article suivant, consacré à une danseuse : http://www.dailymail.co.uk/news/article-3653215/Schoolgirllegamputated-knee-foot-attached-stump-suffering-rare-bone-cancerdefies-odds-competitive-ballet-dancer.html. Cet article laisse imaginer la quantité de travail nécessaire pour faire fonctionner ces différentes technologies.

Envisager des solutions logicielles

Aujourd'hui, parmi les utilisateurs d'ordinateurs, nombreux sont ceux qui utilisent une solution logicielle pour répondre à des besoins particuliers. Un exemple fameux est celui du lecteur d'écran Job Access With Speech (JAWS) (http://www.freedomscientific.com/Products/

Blindness/JAWS) qui transcrit par synthèse vocale le contenu de l'écran à l'aide de méthodes sophistiquées. Comme on peut l'imaginer, toutes les techniques sur lesquelles reposent la science des données et l'IA pour conditionner les données, les interpréter, puis produire un résultat sont sollicitées dans le logiciel JAWS, ce qui en fait un bon moyen d'apprécier les capacités et les limites des solutions logicielles. Si vous voulez savoir ce que cette solution pourrait vous apporter, le mieux est de télécharger et d'installer le logiciel, puis de vous en servir les yeux bandés pour exécuter certaines tâches sur votre système (évitez cependant tout ce qui pourrait vous effrayer, car cela vous ferait faire des erreurs).



Les logiciels d'accessibilité permettent aux utilisateurs ayant des besoins particuliers de réaliser des tâches incroyables. Ils permettent aussi de comprendre ce que c'est que d'avoir des besoins particuliers. Il existe un nombre considérable d'applications de ce genre, mais un bon exemple est celui de Vischeck, à l'adresse http://www.visualisingdata.com/resources/vischeck/, qui vous fait voir des images comme les voient les gens qui sont daltoniens. Naturellement, vous constaterez que ces personnes voient tout de même des couleurs. Simplement, ils les voient comme nous voyons d'autres couleurs.

Dépendre d'un accroissement des capacités matérielles

Pour un certain nombre de besoins particuliers, un simple logiciel ne suffit pas. La section « Envisager l'utilisation d'exosquelettes », précédemment dans ce chapitre, vous donne une idée de la manière dont on utilise actuellement des exosquelettes pour prévenir les blessures, pour accroître les capacités humaines naturelles ou pour

répondre à certains besoins spécifiques (comme permettre à un paraplégique de marcher). Cependant, d'autres types d'accroissement des capacités matérielles sont destinés à d'autres besoins, et dans une vaste majorité des cas, un certain niveau d'IA est nécessaire pour que le système puisse fonctionner correctement.

Prenons l'exemple des systèmes de contrôle par le regard (http://www.eyegaze.com/). Les premiers systèmes utilisaient un tableau de lettres installé au-dessus du moniteur. Le quadriplégique regardait une lettre en particulier, et la direction de son regard était captée par deux caméras (placées de part et d'autre du moniteur) et transmise à l'ordinateur. Un quadriplégique pouvait ainsi faire exécuter des tâches simples par l'ordinateur.

Dans certains systèmes de pilotage oculaire, un bras robotique relié à l'ordinateur exécutait des tâches simples mais importantes, comme tendre un verre à l'utilisateur ou lui gratter le nez. Dans les systèmes plus récents, on peut relier directement le cerveau de l'utilisateur au bras robotique, si bien que l'utilisateur peut, par exemple, manger sans l'aide de personne (voir https://www.engadget.com/2017/03/29/paralyzed-man-first-to-move-his-arm-by-thinking-about-it/).

Voir l'IA faire fonctionner des prothèses

On peut trouver de nombreux exemples d'utilisation de l'IA pour les prothèses. Certes, les prothèses passives existent encore, mais aujourd'hui on privilégie généralement les approches dynamiques dans lesquelles la prothèse fonctionne à l'aide d'une IA. Un des exemples les plus remarquables est le pied bionique mis au point

par Hugh Herr (https://www.smithsonianmag.com/innovation/future-robotic-legs-180953040/), qui fonctionne si bien qu'il lui permet même de pratiquer l'escalade. Vous pouvez voir une démonstration de Hugh Herr sur la chaîne TED à l'adresse https://www.youtube.com/watch?v=CDsNZJTWw0w.



Il se peut qu'un dilemme moral se pose dans l'avenir (mais pas pour le moment, heureusement), lorsque les prothèses permettront à ceux qui les portent de surpasser de façon substantielle les capacités humaines. Dans le film *Eon Flux*, par exemple, Sithandra possède des mains à la place des pieds. Il s'agit d'une sorte de prothèse greffée à une personne qui avait auparavant des pieds normaux. La question qui se pose est de savoir si ce genre de prothèse est valable, utile, ou même souhaitable. À un moment donné, il sera nécessaire de décider dans quelle mesure l'utilisation d'une prothèse ne doit pas faire d'un être humain autre chose qu'un être humain (en supposant que nous voulions rester des humains plutôt qu'évoluer vers un autre stade). Évidemment, on n'imagine pas aujourd'hui que quelqu'un puisse avoir des mains en guise de pieds.

Exploiter de nouvelles méthodes d'analyse et de diagnostic

En médecine comme ailleurs, c'est surtout dans l'analyse des données que l'IA excelle. Certains sites Internet sont entièrement consacrés au rôle joué par l'IA dans la médecine moderne, par exemple celui-ci : http://medicalfuturist.com/category/blog/digitalized-care/artificial-intelligence/.

On pourrait penser qu'il suffit à un spécialiste d'obtenir un cliché d'un éventuel site tumoral pour être en mesure de délivrer un diagnostic fiable. Cependant, les techniques qui permettent d'obtenir ce cliché consistent généralement à traverser des tissus qui ne font pas partie du site tumoral, si bien que le résultat est obscurci. Par ailleurs, un médecin qui examine une tumeur dans ses premiers stades désire obtenir la meilleure information possible.

Non seulement l'IA est une aide au diagnostic pour identifier les tumeurs quand elles sont petites, et avec plus de précision, mais elle accélère aussi considérablement le processus d'analyse. Pour beaucoup de maladies, le temps est un facteur essentiel.

Si les progrès que permet l'IA dans ce domaine en termes de détection et de rapidité sont remarquables, ce qui fait vraiment la différence, c'est d'abord la possibilité d'utiliser l'IA de différentes manières pour compiler les données dans l'Internet des objets. Quand l'IA détecte une maladie chez un patient, elle peut vérifier automatiquement ses antécédents et afficher à l'écran les informations utiles en même temps que les images permettant le diagnostic. Le médecin dispose désormais de toutes les informations pertinentes concernant le patient lorsqu'il doit formuler un diagnostic et envisager une procédure particulière.

Concevoir de nouvelles techniques chirurgicales

Aujourd'hui, on utilise couramment les robots et l'IA dans des procédures chirurgicales. Certaines opérations seraient même pratiquement impossibles sans cela. L'utilisation de cette technologie a cependant commencé il y a peu de temps, sachant que le premier robot utilisé en chirurgie, Arthrobot, a fait son apparition

en 1983 seulement. Elle a permis de réduire le risque d'erreur, d'améliorer les résultats, de raccourcir le temps de guérison, et de façon générale, de rendre la chirurgie moins onéreuse à long terme. Les sections qui suivent décrivent l'utilisation des robots et de l'IA dans divers aspects de la chirurgie.

Formuler des suggestions

Il y a plusieurs façons d'envisager les suggestions dans le domaine chirurgical. Ainsi, par exemple, une IA peut analyser toutes les données relatives à un patient et suggérer au chirurgien la meilleure approche à retenir, compte tenu des antécédents du patient en question. Si ce travail était effectué par le chirurgien, cela prendrait davantage de temps et le risque d'erreur serait plus élevé. L'IA ne se fatigue pas et ne se déconcentre pas. Elle prend systématiquement en compte toutes les données, toujours de la même manière.

Malheureusement, même avec l'aide d'une IA, des surprises peuvent se produire au cours d'une opération chirurgicale. C'est là qu'intervient le niveau de suggestion suivant. D'après l'article de la page https://www.huffingtonpost.com/entry/the-role-of-ai-in-surgery_us_58d40b7fe4b002482d6e6f59, les médecins peuvent maintenant disposer d'un appareil qui fonctionne de la même manière qu'Alexa, Siri et Cortana, ces assistants personnels intelligents (vous pourriez en avoir un chez vous). Non, cet appareil n'est pas destiné à diffuser pendant l'opération une musique choisie par le chirurgien, mais celui-ci peut s'en servir pour localiser certaines informations sans être obligé d'interrompre son acte. Le patient profite ainsi de l'équivalent d'un second avis et évite de subir des complications imprévues. En fait, l'appareil ne fait rien d'autre que reproduire une recherche

déjà prévue par d'autres médecins et disponible sur demande, pour les besoins de l'opération : il ne s'agit pas d'une véritable réflexion.

Pour qu'une opération puisse avoir lieu dans de bonnes conditions, il faut aussi analyser toutes les prises de vue que les médecins tiennent à obtenir. Ici, la rapidité est un avantage qu'offre l'IA par rapport au radiologue. Un produit comme Enlitic (https://www.enlitic.com/), une technologie d'apprentissage profond, est capable d'analyser des radiographies en l'espace de quelques millisecondes, c'est-à-dire dix mille fois plus vite qu'un radiologue. En outre, la performance de ce système est de 50 % supérieure pour ce qui est de catégoriser les tumeurs, et le taux de faux négatifs est plus faible, de l'ordre de 0 % contre 7 % chez une personne humaine. Un autre produit de la même catégorie, Arterys (https://arterys.com/), effectue une scintigraphie cardiaque en 6 à 10 minutes, contre une heure habituellement. Par ailleurs, le patient n'est plus obligé de retenir sa respiration. Ce qui est formidable également, c'est que ce système produit des données en plusieurs dimensions : anatomie du cœur en 3D, débit sanguin et direction du débit sanguin, tout cela en si peu de temps. Vous pouvez voir une vidéo sur Arterys à l'adresse https://www.youtube.com/watch?v=IcooATgPYXc.

Assister un chirurgien

La plupart du temps, la robotique ne remplace pas le chirurgien mais elle l'assiste. Le premier robot de chirurgie, le système PUMA, est apparu en 1986. Il servait à pratiquer une biopsie neurochirurgicale extrêmement délicate. Il s'agissait d'un type de chirurgie non laparoscopique. La chirurgie laparoscopique est minimalement invasive : on pratique simplement un ou

plusieurs petits trous pour accéder à un organe, par exemple la vésicule biliaire, afin de le retirer ou de le réparer. Les premiers robots n'étaient pas assez habiles pour exécuter cette tâche.

TRAVAILLER DANS DES PAYS DU TIERS-MONDE

On se dit souvent qu'aucune de ces formidables technologies qu'utilisent aujourd'hui les professionnels de la santé ne bénéficie aux populations des pays du tiers-monde. En réalité, certaines de ces technologies, comme les produits de Bay Labs (https://baylabs.io/), sont spécifiquement destinées à ces pays. Des médecins ont utilisé ces technologies en Afrique, notamment au Kenya pour identifier les symptômes de cardiopathie rhumatismale chez les enfants. Au cours d'une visite effectuée en septembre 2016, des médecins se sont servis du matériel Bay Labs pour examiner 1 200 enfants en l'espace de quatre jours, ce qui leur a permis de diagnostiquer une cardiopathie rhumatismale ou une maladie cardiaque congénitale chez 48 d'entre eux. Sans l'IA, ce matériel n'aurait pas pu exister, sachant qu'il n'aurait jamais été assez petit ni assez facile à utiliser pour qu'il soit possible de s'en servir dans un tel contexte.

Depuis 2000, le système chirurgical da Vinci permet de pratiquer des actes de chirurgie laparoscopique robotisée à l'aide d'un système optique 3D. C'est le praticien qui dirige les mouvements du robot, mais c'est le robot qui effectue l'opération. Le chirurgien a les yeux fixés sur un écran haute définition qui lui permet de mieux suivre l'opération que s'il opérait lui-même. Par ailleurs, avec le système da Vinci, les trous pratiqués sont plus petits, et cela réduit les risques d'infection.

Cependant, l'aspect le plus important est le fait que ce système accroisse les capacités du praticien. Ainsi, par

exemple, le système da Vinci élimine le problème des éventuels tremblements de la main du médecin, un peu comme la fonction de stabilisation d'image dans un appareil photo numérique. Le système atténue aussi les vibrations, et les mouvements sont extrêmement précis, bien plus précis que ceux d'une personne humaine.



Le système chirurgical da Vinci est complexe et extrêmement flexible. Aux États-Unis, la FDA l'a homologué pour les actes chirurgicaux suivants, aussi bien chez les enfants que chez les adultes :

- chirurgie urologique ;
- chirurgie générale laparoscopique ;
- chirurgie thoracoscopique générale non cardiovasculaire ;
- procédures de cardiotomie assistée par thoracoscopie.

Par-delà tout ce jargon médical, ce qu'il convient de retenir, c'est que le système chirurgical da Vinci peut exécuter un certain nombre de tâches sans l'intervention directe du chirurgien. À un moment donné, les robots de chirurgie seront plus autonomes, si bien que le praticien humain restera davantage à distance du patient durant l'opération. Un jour, plus personne ne pénétrera dans le bloc opératoire avec le patient, ce qui éliminera quasiment les risques d'infection. Pour plus de détails sur le système da Vinci, consultez la page http://www.davincisurgery.com/da-vinci-surgery/da-vinci-surgical-system/.

Remplacer le chirurgien

Dans *La Guerre des étoiles*, on voit des robots chirurgiens qui rafistolent des humains. C'est même à se demander

s'il y a des médecins humains. Théoriquement, certaines opérations chirurgicales pourraient être pratiquées par des robots dans l'avenir, mais cette possibilité reste lointaine. Par rapport aux applications industrielles actuelles, d'importants progrès seraient nécessaires en matière de robotique. Les robots actuels ne sont pas très autonomes, et des interventions humaines restent indispensables pour leur installation et leur paramétrage.

Cependant, des avancées ont lieu dans le domaine de la robotique appliquée à la chirurgie. Le robot Smart Tissue Autonomous Robot (STAR), par exemple, sait recoudre l'intestin d'un porc mieux que les chirurgiens, comme on peut le voir sur la page https://www.france24.com/fr/20160505-premiere-fois-robot-chirurgien-effectue-seul-une-operation-tissus-mous. STAR a opéré sous la supervision d'une équipe de médecins, mais il a exécuté la tâche tout seul, et c'est un pas en avant considérable dans le domaine de la chirurgie robotisée. La vidéo à l'adresse https://www.youtube.com/watch?v=vb79-_hGLkc montre bien comment évolue la chirurgie aujourd'hui.

EXÉCUTER DES TÂCHES EN RECOURANT À L'AUTOMATION

L'IA fait des merveilles dans le domaine de l'automation. Elle ne s'écarte jamais de la procédure, elle n'est jamais fatiguée et elle ne commet jamais d'erreurs, dans la mesure où la procédure initiale est correcte. Contrairement aux humains, l'IA n'a jamais besoin de vacances ni d'un week-end de repos, ni même d'une réduction du temps de travail (il est vrai que dans la profession médicale, nombreux sont ceux qui ne peuvent pas envisager tout cela). En conséquence, la même IA peut s'occuper d'un patient au moment du petit déjeuner, puis du déjeuner et du dîner. L'IA présente donc d'entrée de jeu d'importants avantages, si l'on se fonde sur des critères de cohérence, de précision et de durée (pour savoir dans quel domaine l'IA ne fait pas l'affaire, voir l'encadré « Biais, sympathie et empathie »). Les sections qui suivent traitent des différentes manières dont l'IA peut permettre une automatisation par un meilleur accès aux ressources, notamment aux données.

Travailler avec des dossiers médicaux

C'est avant tout dans le domaine des dossiers médicaux qu'une IA peut améliorer la pratique de la médecine. Dans le passé, on enregistrait toujours les données des patients dans des dossiers sur support papier. Dans les hôpitaux, le personnel médical pouvait aussi utiliser une ardoise pour noter les données quotidiennes de chaque patient. Les médecins utilisent souvent des graphiques, et ils peuvent aussi utiliser des notes. Avec toutes ces sources d'information dans tant d'endroits différents, il était difficile de suivre le patient de façon vraiment efficace. Quand une IA est utilisée conjointement avec une base de données

informatisée, l'information est accessible, cohérente et fiable. Des produits comme Deepmind Health (https://deepmind.com/applied/deepmind-health/working-partners/health-research-tomorrow/), de Google, permettent au personnel de rassembler les informations sur les patients et de détecter des tendances dans les données qui ne sont pas évidentes.



Les médecins n'utilisent pas nécessairement les dossiers de la même manière qu'on les utilise dans d'autres métiers. Des produits comme WatsonPaths (http://www.research.ibm.com/cognitive-computing/watson/watsonpaths.shtml), d'IBM, leur permettent d'interagir d'une nouvelle manière avec toutes sortes de données sur les patients afin de formuler de meilleurs diagnostics et de prendre de meilleures décisions concernant les soins à dispenser. Pour voir comment ce produit fonctionne, regardez cette vidéo : https://www.youtube.com/watch?v=07XPEqkHJ6U.

La médecine est un travail d'équipe qui réunit des personnes exerçant des spécialités différentes. Cependant, en se penchant sur le processus, on ne tarde pas à se rendre compte que ces personnes ne communiquent pas suffisamment entre elles, parce qu'elles sont trop occupées à soigner les patients. Un produit comme CloudMedX (http://www.cloudmedxhealth.com/) reçoit les données de toutes les parties en présence et effectue une analyse de risque. Ce logiciel permet ainsi de localiser les domaines potentiellement problématiques qui sont susceptibles d'affecter l'état du patient. En d'autres termes, ce produit se charge de la communication qui existerait entre les parties en présence si elles n'étaient pas submergées de travail.

Prédire l'avenir

Un bon exemple de logiciel prédictif exploitant les dossiers médicaux est CareSkore (https://www.careskore.com/), qui utilise des algorithmes pour déterminer la probabilité qu'un patient ait besoin d'être à nouveau hospitalisé. Le personnel hospitalier peut examiner les raisons d'une éventuelle réadmission et résoudre le problème avant que le patient quitte l'hôpital, de telle sorte que sa réadmission risque moins d'être nécessaire. En parallèle, Zephyr Health (https://zephyrhealth.com/) aide le médecin à évaluer différentes thérapies et à choisir celles qui ont le plus de chances d'aboutir à un résultat positif : là encore, on réduit ainsi le risque que le patient ait besoin d'être hospitalisé à nouveau. Pour plus de détails sur Zephyr Health, regardez cette vidéo : https://www.youtube.com/watch ? v=9y930hioWjw.

Sous certains aspects, vos gènes sont une carte indiquant ce qui vous arrivera dans l'avenir. La connaissance de vos données génétiques vous permet donc de mieux comprendre quels sont vos points forts et vos points faibles et ainsi, de vivre mieux. Deep Genomics (https://www.deepgenomics.com/) est en train de découvrir l'impact que des mutations de vos gènes auront sur votre personne. Les mutations ne produisent pas nécessairement un résultat négatif : certaines mutations sont bénéfiques, et mieux connaître ce domaine peut être aussi une expérience positive. Pour plus de détails, regardez la vidéo suivante : https://www.youtube.com/watch ? v=hVibPJyf-xg.

Rendre les procédures plus sûres

Les médecins ont besoin de disposer d'un grand nombre de données pour pouvoir prendre les bonnes décisions. Or, les données étant éparpillées, les médecins ne sont pas toujours en mesure de les analyser assez vite pour pouvoir prendre les bonnes décisions. Pour rendre les procédures plus sûres, un médecin a besoin non seulement d'accéder aux données, mais aussi de disposer d'un moyen de les organiser et de les analyser qui soit adapté à sa spécialité. Un produit qui rend ce service est Oncora Medical (https://oncoramedical.com/), qui collecte et organise les données médicales pour les radio-oncologues. Cela leur permet de diriger la bonne quantité de rayonnements là où il faut et d'obtenir ainsi un meilleur résultat avec moins de risque d'effets secondaires imprévus.

Si les médecins ont des difficultés à obtenir les informations nécessaires, c'est aussi parce que les machines qu'ils utilisent sont généralement coûteuses et volumineuses. Un innovateur, Jonathan Rothberg, a décidé de changer tout cela grâce au Butterfly Network (https://www.butterflynetwork.com/#News). Il s'agit d'un appareil de la taille d'un iPhone et capable d'effectuer une IRM et une échographie. L'image sur le site Internet n'est rien moins qu'étonnante.

Créer de meilleurs médicaments

Aujourd'hui, tout le monde se plaint du prix des médicaments. La médecine peut accomplir des miracles, mais les traitements sont si coûteux que certains finissent

par hypothéquer leur maison pour les obtenir. Ce problème est dû en partie au temps nécessaire pour tester les médicaments. Une analyse des tissus pour observer les effets d'un nouveau médicament peut demander un an. Heureusement, un produit comme 3Scan (http://www.3scan.com/) permet d'obtenir la même analyse des tissus dans un délai considérablement plus court, parfois en l'espace d'une journée seulement.

Bien sûr, ce qui serait mieux encore, c'est que les laboratoires pharmaceutiques aient une idée plus précise des produits qui auront le plus de chances d'être efficaces et de ceux qui ne le seront pas avant de commencer à investir de l'argent dans la recherche. Atomwise (http://www.atomwise.com/) utilise une énorme base de données de structures moléculaires pour effectuer des analyses en vue de savoir quelles molécules répondront à un besoin particulier. En 2015, des chercheurs ont utilisé Atomwise pour mettre au point des médications qui rendraient le virus Ebola moins susceptible de contaminer l'entourage. Pour effectuer une analyse qui aurait demandé à des chercheurs plusieurs mois, ou peut-être plusieurs années, Atomwise n'a mis qu'une journée. Imaginons ce scénario au milieu d'une épidémie potentiellement mondiale. Si Atomwise peut effectuer en un jour l'analyse requise pour rendre la bactérie ou le virus non contagieux, l'épidémie peut alors être jugulée avant de se répandre.

Par ailleurs, les laboratoires produisent de grandes quantités de médicaments. La raison de cette productivité impressionnante, en dehors du profit, est que chaque individu est un petit peu différent. Un médicament qui est efficace et ne provoque aucun effet secondaire chez une personne peut ne pas être efficace du tout chez une autre personne, et il peut même lui faire du mal. Turbine (http://turbine.ai/) permet aux laboratoires de réaliser des simulations pour déterminer les médicaments

les plus susceptibles d'agir efficacement chez une personne particulière. Turbine privilégie actuellement les traitements contre le cancer, mais il est facile de voir comment la même approche peut être pertinente dans un certain nombre d'autres domaines.



Une médication peut prendre des formes variées. Il n'y a pas que les comprimés et les médicaments injectables. Notre organisme produit une vaste gamme de traitements sous forme de microbiomes. Il contient dix fois plus de microbes que de cellules, et un certain nombre de ces microbes nous sont indispensables : sans eux, nous serions vite morts. Whole Biome (https://www.wholebiome.com/) utilise une variété de méthodes pour améliorer le fonctionnement de ces microbiomes, de telle sorte que nous n'ayons pas nécessairement besoin d'un traitement par comprimés ou autres modes d'administration. Pour plus de précisions, regardez la vidéo suivante : https://www.youtube.com/watch?v=t1Y2AckssyI.

Certaines compagnies n'ont pas encore réalisé leur potentiel, mais ils le réaliseront sans doute un jour ou l'autre. C'est le cas de Recursion Pharmaceuticals (https://www.recursionpharma.com/), qui utilise l'automation pour explorer des façons de résoudre de nouveaux problèmes en utilisant des médicaments connus, des produits bioactifs et des produits pharmaceutiques n'ayant pas fait leurs preuves. Cette compagnie a connu un certain succès dans la guérison de maladies génétiques rares, et elle s'est fixé comme objectif de guérir 100 maladies au cours des dix ans à venir (elle a placé la barre très haut).

Combiner les robots et les professionnels de la santé

Des robots semi-autonomes à capacités limitées commencent à s'intégrer dans la société. Au Japon, ces robots sont utilisés depuis déjà un bon moment (voir https://www.japantimes.co.jp/news/2017/05/18/national/science-health/japans-nursing-facilities-using-humanoid-robots-improve-lives-safety-el-derly/). Ces robots font aussi leur apparition en Amérique sous la forme de Rudy (voir http://www.roboticstrends.com/article/rudy_assistive_robot_helps_elderly_age_in_place/health_sports). Dans la plupart des cas, ces robots peuvent exécuter des tâches simples, comme rappeler à quelqu'un qu'il doit prendre ses médicaments ou jouer à des jeux simples, sans qu'une intervention humaine soit souvent nécessaire. Cependant, le cas échéant, un médecin ou autre professionnel de santé peut prendre le contrôle du robot depuis un endroit distant et s'en servir pour effectuer des tâches plus complexes. Ainsi, le praticien obtient une aide instantanée lorsqu'il en a besoin, ce qui réduit le risque de préjudice pour le patient et permet de limiter les coûts.



Ces robots n'en sont encore qu'à leurs balbutiements, mais on peut s'attendre à les voir évoluer. S'ils ne sont que des outils d'assistance au personnel médical et ne peuvent pas réellement remplacer un médecin ni une infirmière pour un certain nombre de tâches spécialisées, ils assurent tout de même la surveillance constante dont les patients ont besoin, ainsi qu'une présence réconfortante. En outre, ces robots permettent d'éviter de devoir confier à des personnes humaines des tâches triviales et répétitives qu'ils savent déjà très bien accomplir

(comme distribuer des comprimés, fournir des rappels ou assister le patient dans ses déplacements).

Chapitre 8

Utiliser l'IA pour améliorer l'interaction humaine

DANS CE CHAPITRE

- » **Communiquer autrement**
- » **Partager des idées**
- » **Jouer la carte du multimédia**
- » **Améliorer la perception sensorielle**

Les interactions entre les personnes peuvent prendre des formes très variées. La plupart des gens ne se rendent même pas compte de la variété des moyens par lesquels nous communiquons. Or, la communication, ce n'est pas seulement écrire ou parler. Il existe bien d'autres formes d'interactions, entre autres, le contact visuel, la qualité tonale, et même les odeurs (voir https://www.smithsonianmag.com/science-nature/thetruth-about-pheromones-100363955/). Un exemple d'émulation d'une interaction humaine améliorée est le nez électronique, une technologie qui fait appel à la fois à l'électronique, à la biochimie et à l'intelligence artificielle et qui trouve des applications dans tout un ensemble de domaines de l'industrie et de la recherche (voir https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3274163/). Ce chapitre est plutôt axé sur les formes

classiques de communication, notamment le langage corporel. Vous allez pouvoir mieux comprendre comment l'IA vous permet d'améliorer la communication humaine par des moyens moins onéreux que la fabrication de votre propre nez électronique.

L'IA peut aussi faciliter les échanges d'idées. Dans certains cas, elle inaugure des méthodes de communication entièrement nouvelles, mais bien souvent elle offre une méthode généralement subtile (pas toujours) pour améliorer les moyens actuels d'échanger les idées. Nous avons besoin d'échanger des idées pour pouvoir créer de nouvelles technologies, exploiter les technologies existantes et concevoir les technologies qui nous permettront de faire progresser nos connaissances. Or, compte tenu de leur caractère abstrait, les idées sont parfois particulièrement difficiles à échanger. L'IA peut donc créer les ponts nécessaires entre les gens.

Il fut un temps où c'était généralement l'écriture qui permettait d'enregistrer des connaissances pour les partager avec autrui. Dans certains cas, les auteurs augmentaient leur communication en utilisant divers types d'illustrations graphiques. Cependant, si certains peuvent encore se contenter d'utiliser ces deux formes de médias pour acquérir des connaissances, nombreux sont ceux qui ont besoin d'autre chose, c'est pourquoi des sources en ligne comme YouTube (https://www.youtube.com/) ont maintenant autant de succès. Fait intéressant, vous pouvez augmenter le pouvoir du multimédia, déjà substantiel, en utilisant l'IA, et ce chapitre vous explique comment.

La dernière section de ce chapitre vous permet de comprendre comment une IA peut vous doter d'une perception sensorielle presque surhumaine. Finalement, ce nez électronique, peut-être aimeriez-vous l'acquérir : il présente des avantages significatifs lorsqu'il s'agit de

détecter des odeurs qui sont nettement moins aromatiques que celles que nous sommes capables de sentir. Imaginez que vous soyez doté des mêmes capacités olfactives qu'un chien (qui utilise 100 millions de récepteurs olfactifs, tandis que l'être humain en utilise 1 million). Il s'avère qu'il existe deux moyens d'atteindre cet objectif : un accès indirect au moyen d'un moniteur, et la stimulation directe de la perception sensorielle humaine.

Développer de nouvelles façons de communiquer

La communication à l'aide d'un langage développé entre les humains s'est d'abord faite oralement. Le seul problème avec la communication parlée est que les deux parties doivent être assez proches l'une de l'autre pour pouvoir se parler. La communication écrite est supérieure à la communication parlée par bien des aspects, car elle permet des communications différées entre deux parties qui n'ont pas besoin de se voir. Les trois principales méthodes de communication humaine non verbale reposent sur :

- **Les alphabets** : l'abstraction des composantes des paroles humaines et des symboles.
- **Le langage** : l'enchaînement de mots ou de symboles de manière à créer des phrases et à transmettre des idées sous forme écrite.
- **Le langage corporel** : l'augmentation du langage par le contexte.

Les deux premières méthodes sont des abstractions directes de la parole. Elles ne sont pas toujours faciles à mettre en œuvre, mais elles sont utilisées depuis des

millénaires. Le langage corporel est le volet le plus difficile à mettre en œuvre car il s'agit de créer une abstraction d'un processus physique. L'écriture permet de véhiculer le langage corporel à l'aide d'une terminologie spécifique, comme celle présentée sur la page https://writerswrite.co.za/cheat-sheets-for-writing-body-language/. Cependant, la transcription de la parole ne suffit pas, c'est pourquoi elle est augmentée à l'aide de symboles, comme les émoticônes et les *emoji*. Les sections qui suivent traitent de ces sujets plus en détail.

Créer de nouveaux alphabets

Dans son introduction, cette section présente deux nouveaux alphabets utilisés à l'âge de l'informatique : les émoticônes et les emoji (https://emojipedia.org/). Les sites Internet sur lesquels vous trouverez ces deux types d'alphabets en présentent souvent des centaines. Nous sommes capables d'interpréter ces alphabets iconiques sans trop de difficultés, car ils imitent des expressions faciales, mais comme une application informatique ne possède pas notre sens artistique, l'ordinateur a souvent besoin d'une IA pour savoir quelle émotion une personne essaie de transmettre à l'aide de ces petites images. Heureusement, il existe des listes standardisées, comme celle-ci : https://unicode.org/emoji/charts/full-emoji-list.html. Naturellement, une liste standardisée n'aide pas réellement à traduire. L'article de la page https://www.geek.com/tech/ai-trained-on-emoji-can-detectsocial-media-sarcasm-1711313/ fournit plus de détails sur la manière dont il est possible d'apprendre à une IA à interpréter un emoji (et par extension, des émoticônes) et à y répondre. On peut

voir fonctionner un exemple de ce processus sur la page https://deepmoji.mit.edu/.

L'émoticône est une technologie plus ancienne, que beaucoup de gens cherchent à oublier (mais ils n'y parviendront probablement pas). Quant à l'emoji, il est assez nouveau et assez amusant pour faire l'objet d'un film (voir https://www.amazon.com/exec/obidos/ASIN/B0746ZZR71/datacservip0f-20/). Vous pouvez aussi vous servir de l'IA de Google pour convertir vos *selfies* en emoji (voir https://www.fastcodesign.com/90124964/exclusivenew-google-tool-uses-ai-to-create-custom-emoji-of-you-from-a-selfie). Et si vous ne souhaitez vraiment pas passer en revue les 2666 emoji officiels qu'accepte Unicode (et encore moins les 564 quadrillions d'emoji que l'application Allo, de Google, à l'adresse https://allo.google.com/, est capable de générer), vous pouvez demander à Dango (https://play.google.com/store/apps/details?id=co.dango.emoji.gif&hl=fr) de vous suggérer un emoji approprié (voir https://www.technologyreview.com/s/601758/this-appknows-just-the-right-emoji-for-any-occasion/).



Depuis le début de l'écrit, les hommes créent de nouveaux alphabets pour répondre à des besoins spécifiques. Parmi ces alphabets, les émoticônes et les emoji sont liés à l'utilisation de l'Internet et de l'IA. Une IA peut même être nécessaire pour pouvoir les gérer tous.

Automatiser la traduction

L'absence d'une langue commune a toujours posé un problème dans le monde. Certes, l'anglais est devenu la

langue internationale, mais elle n’est pas complètement universelle. Sachant qu’une traduction peut être onéreuse, fastidieuse et source d’erreurs, les traducteurs, même s’ils sont souvent indispensables, ne sont pas toujours la meilleure solution. Pour quiconque ne bénéficie pas de l’assistance d’un traducteur, être confronté à un texte en langue étrangère peut poser des problèmes, d’où l’intérêt que représentent des applications comme Google Traduction (voir Figure 8.1).

Une chose qu’il faut remarquer sur la Figure 8.1 est que Google Traduction vous propose de détecter automatiquement la langue à votre place. Cette fonctionnalité est la plupart du temps très efficace, en partie grâce au système Google Neural Machine Translation (GNMT). Ce système est capable d’analyser une phrase entière et d’effectuer de meilleures traductions que les applications qui créent la traduction en traitant des mots ou des groupes de mots (voir http://www.wired.co.uk/article/google-ai-language-create).



Plus impressionnant encore, GNMT peut traduire d’une langue à une autre même en l’absence d’un traducteur spécifique, grâce à un langage artificiel, un *interlingua* (voir https://en.oxforddictionaries.com/definition/interlingua). Cependant, il est important de se rendre compte qu’un interlingua ne fonctionne pas comme un traducteur universel : c’est plutôt un pont universel. GNMT ne sait pas traduire du chinois à l’espagnol, par exemple, mais il peut traduire du chinois à l’anglais et de l’anglais à l’espagnol. Grâce à la construction d’un réseau en 3D représentant ces trois langues (l’interlingua), GNMT est capable de créer sa propre traduction du chinois à l’espagnol. Malheureusement, ce système ne permettrait pas de traduire du chinois au martien, car il n’existe encore aucune méthode permettant de comprendre et

traduire du martien vers une langue humaine. Pour que GNMT puisse effectuer une telle tâche, il faudrait avoir mis au point une traduction de base.

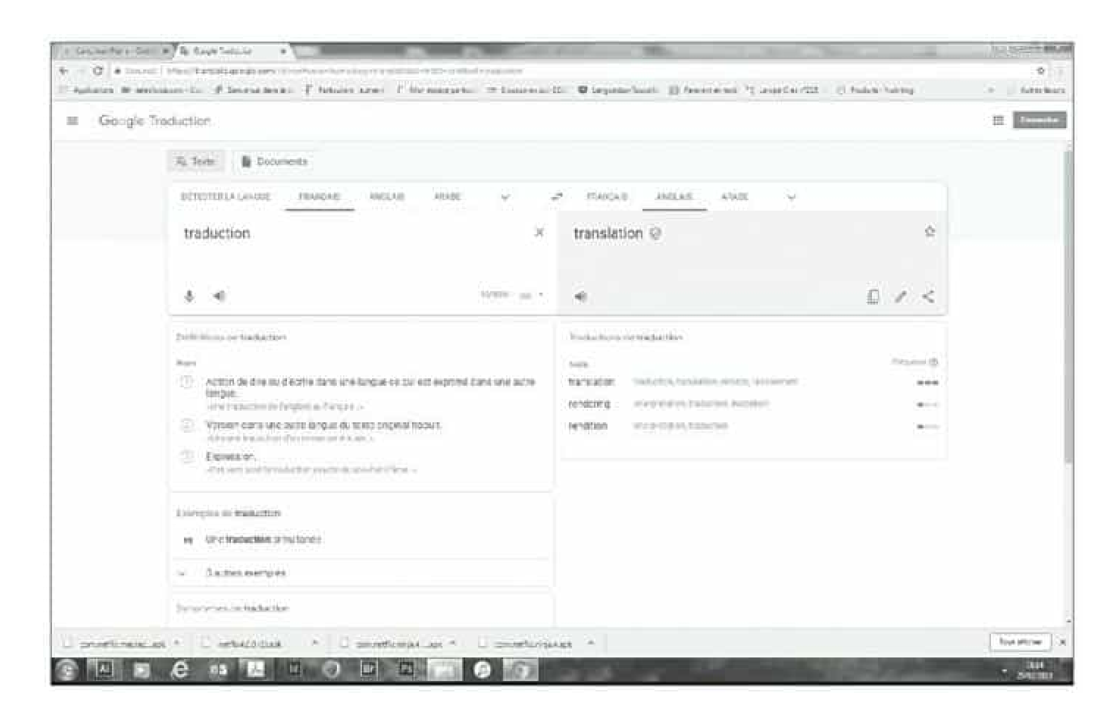


FIGURE 8.1 : Google Traduction est un exemple d'IA exécutant une tâche quotidienne essentielle.

Intégrer le langage corporel

Une partie importante de la communication humaine se fait par le langage corporel, et c'est la raison pour laquelle les émoticônes et les emoji sont importants. Cependant, les gens sont de plus en plus habitués à travailler directement avec des caméras pour créer des vidéos et autres formes de communication sans écriture. Dans ce cas, l'ordinateur pourrait recevoir l'input sonore de la voix humaine, l'analyser et la traduire en signes représentant le discours humain, puis traiter ces signes pour répondre à une requête, de façon similaire au fonctionnement d'Alexa, Google Home et autres assistants vocaux.



Malheureusement, la simple traduction sous forme de signes du langage parlé ne fait pas l'affaire, car tout le problème de la communication non verbale reste posé. Il faut alors que l'IA soit capable de lire directement le langage corporel. L'article de la page https://www.cmu.edu/news/stories/archives/2017/july/computer-reads-body-language.html traite de certains problèmes que les développeurs doivent résoudre pour pouvoir rendre possible la lecture du langage corporel. L'image qui figure au début de l'article donne une idée de la manière dont la caméra de l'ordinateur doit saisir les positions de la personne pour pouvoir lire le langage corporel, et l'IA a souvent besoin de recevoir l'input de plusieurs caméras, sachant que dans le champ d'une seule caméra, une partie de l'anatomie humaine est cachée. La lecture du langage corporel consiste à interpréter les caractéristiques humaines suivantes :

- la posture ;
- les mouvements de la tête ;
- l'expression faciale ;
- le contact visuel ;
- la gestuelle.

EN FINIR AVEC LE MYTHE DES LANGAGES SPÉCIFIQUES À L'IA

Vous lirez peut-être des articles comme celui de la page https://www.fastcodesign.com/90132632/ai-is-inventing-its-own-perfect-languages-should-we-let-it qui semble indiquer que certaines IA seraient capables de créer de nouveaux langages qui seraient incompréhensibles pour les humains, puis de les utiliser pour communiquer. Cette idée suggère un scénario à la Terminator (https://www.amazon.com/exec/obidos/ASIN/B00938UVC2/datacservip0f-20/), et comme le film en question, ce n'est que de la fiction. En lisant bien l'article, vous constaterez cependant que le langage n'est pas nouveau et que la communication est étonnamment aléatoire. Dans les chapitres précédents, nous avons étudié le concept de compréhension en IA. En réalité, une IA ne comprend pas un langage : elle le convertit en série d'unités lexicales qu'elle analyse sous forme de représentations mathématiques de mots qui apparaissent dans une table de consultation. La section « Analyser l'argument de la chambre chinoise » du Chapitre 5 traite de l'utilisation de tables de consultation pour donner l'impression qu'un système est intelligent et comprend un langage, alors qu'en réalité il ne le comprend pas.

Bien sûr, ces caractéristiques ne sont pas les seules, mais une IA qui peut les interpréter toutes les cinq est déjà capable de fournir une interprétation correcte du langage corporel. Outre le langage corporel, les applications actuelles de l'IA prennent aussi en compte des caractéristiques comme la qualité tonale. On obtient ainsi une IA extrêmement complexe, mais qui reste loin d'approcher ce que le cerveau humain semble pouvoir faire sans effort.



Une IA capable de lire le langage corporel doit aussi pouvoir produire le résultat de cette lecture quand elle interagit avec les humains. Sachant que ce type de lecture n'en est qu'à ses balbutiements, la présentation robotique ou graphique du langage corporel en est à un stade encore moins avancé. Il existe des situations dans lesquelles les robots peuvent aujourd'hui interpréter le langage corporel et répondre de manière appropriée. Comme les robots ne sont pas encore capables de produire des expressions faciales satisfaisantes, d'après l'article de la page http://theconversation.com/realistic-robot-faces-arent-enough-we-needemotion-to-put-us-at-ease-with-androids-43372, dans le meilleur des cas ils imitent la posture, les mouvements de la tête et la gestuelle du langage corporel. Cependant, le résultat n'a rien d'impressionnant.

Échanger des idées

Une IA n'a pas d'idées, car il lui manque à la fois l'intelligence intrapersonnelle et la capacité de comprendre. Une IA peut cependant permettre aux humains d'échanger des idées dans le cadre d'un tout qui soit plus que la somme de ses parties. Dans bien des cas, l'IA n'opère aucun échange d'aucune sorte. Ce sont les humains impliqués dans le processus qui effectuent les échanges en se servant de l'IA pour augmenter le processus de communication. Les sections qui suivent fournissent davantage de détails sur ce processus.

Créer des liens

Un échange d'idées entre deux êtres humains suppose qu'ils se connaissent. Le problème est que souvent les

spécialistes d'un domaine particulier ne se connaissent pas, ou du moins, pas suffisamment pour pouvoir communiquer. Une IA peut effectuer une recherche d'après un flux d'idées fourni par un utilisateur humain, puis créer des liens avec d'autres humains qui auront des idées similaires.

Ce type de création de communication s'observe sur des réseaux sociaux comme LinkedIn (https://www.linkedin.com/), dont le principe est de créer des liens entre les gens sur la base d'un certain nombre de critères. Le réseau de relations d'une personne devient le moyen par lequel l'IA qui est derrière LinkedIn suggère d'autres relations possibles. En fin de compte, l'objet de ces liens, du point de vue de l'utilisateur, est d'avoir accès à de nouvelles ressources humaines, de créer des contacts commerciaux, de décrocher une vente ou d'accomplir d'autres tâches rendues possibles par LinkedIn grâce à ces divers liens.

Augmenter la communication

Deux humains qui veulent échanger des idées ont besoin de pouvoir bien communiquer. Or, les gens, parfois, ne communiquent pas bien, voire pas du tout. Ce n'est pas seulement un problème de traduction, c'est aussi un problème d'idées. Les préjugés sociétaux et personnels peuvent faire obstacle à la communication, et une idée qui est partagée par un groupe peut ne pas être du tout à la portée d'un autre groupe. Par exemple, le point de vue d'un individu sur les lois d'un pays pourra être très différent du point de vue d'un autre individu sur les lois d'un autre pays.

Théoriquement, une IA peut permettre la communication entre des groupes disparates de différentes manières. Bien sûr, la traduction (en supposant qu'elle sera

pertinente) fait partie de ces méthodes. Cependant, au moyen d'un filtrage préalable, une IA peut fournir des indices pour comprendre ce qui est ou n'est pas culturellement acceptable. Grâce à une catégorisation, une IA peut aussi proposer des aides, par exemple d'autres représentations graphiques, afin de permettre que s'établisse une communication utile aux deux parties.

Définir des tendances

Les humains fondent souvent leurs idées sur des tendances. Or, pour que vos interlocuteurs puissent visualiser l'application d'une idée que vous proposez, il faut qu'ils comprennent aussi quelles sont ces tendances, et l'on sait que communiquer ce genre d'information est difficile. L'IA peut effectuer l'analyse des données à plusieurs niveaux et présenter le résultat sous forme graphique. Elle peut analyser les données de façon plus variée et plus rapide qu'un être humain et faire dire aux données ce que l'on a besoin de savoir. Les données sont les mêmes, ce sont leur présentation et leur interprétation qui changent.

Des études montrent que les résultats sont plus parlants sous forme graphique que sous forme de tableaux, et il est certain qu'une tendance se distingue mieux sur un graphique. Comme l'explique la page http://sphweb.bumc.bu.edu/otlt/mph-modules/bs/datapresentation/DataPresentation2.html, les données tabulaires ne sont généralement utilisées que pour présenter des informations particulières, et les graphiques montrent toujours mieux les tendances. Les applications de l'IA permettent aussi de produire plus facilement les résultats graphiques adéquats, pour une demande donnée. Sachant que tout le monde ne voit pas un graphique de la même façon, adapter un graphique à son public est essentiel.

Utiliser le multimédia

La plupart des gens apprennent en sollicitant plusieurs sens, plusieurs approches. Une méthode d'apprentissage qui se révèle efficace pour une personne ne le sera pas pour une autre, qui sera désorientée. Par conséquent, plus on dispose de moyens variés pour communiquer des concepts et des idées, plus on a de chances de pouvoir se faire comprendre. Le multimédia associe généralement le son, l'image, le texte et l'animation, mais il fait parfois davantage encore.

L'IA peut compléter le multimédia de nombreuses manières. Elle peut notamment créer le multimédia. On retrouve l'IA dans des applications de développement et de présentation de médias. Pour traduire les couleurs dans une image, par exemple, une IA peut vous permettre de visualiser les effets de ces changements plus vite que si vous testiez des combinaisons de couleurs une par une (ce qu'on appelle l'approche par la force brute).

Une fois que des idées ont été présentées sous plusieurs formes grâce au multimédia, les destinataires doivent traiter cette information. Une utilisation secondaire de l'IA consiste à traiter l'information par des réseaux de neurones, de différentes manières. La catégorisation du multimédia est une utilisation essentielle de la technologie aujourd'hui. Dans l'avenir, cependant, on peut espérer pouvoir utiliser l'IA pour la reconstitution de scènes en 3D à partir d'images en 2D. Imaginez que les policiers puissent se déplacer sur une scène de crime virtuelle et noter fidèlement tous les détails.

LE MULTIMÉDIA ET LES BESOINS PARTICULIERS

La plupart des gens ont des besoins particuliers, et il est important de prendre en considération les besoins particuliers des gens dans le cadre de leur utilisation du multimédia. L'objet global du multimédia est de communiquer des idées par autant de moyens que possible, de telle sorte que tout le monde puisse comprendre les idées et les concepts que vous voulez présenter. Même quand une présentation est réussie, certaines idées peuvent se perdre si cette présentation n'utilise qu'une seule méthode pour les communiquer. Dans le cas d'une communication exclusivement sonore, par exemple, seules les personnes ayant une bonne oreille recevront le message. Certains ne saisiront cependant pas l'idée, parce qu'ils n'auront perçu que du bruit ou parce qu'ils n'auront pas reçu l'information de la manière requise. Il est essentiel d'utiliser autant de méthodes que possible pour communiquer chaque idée, si l'on veut toucher le plus grand nombre.

Il a été dit que divers types de médias allaient apparaître sous de nouvelles formes. On imaginait, par exemple des journaux qui afficheraient des images dynamiques, comme dans le monde de Harry Potter. Les technologies permettant ce genre de chose existent pour la plupart aujourd'hui, mais leur application est une question de marché. Pour qu'une technologie ait du succès, il faut qu'il y ait un marché correspondant : il faut que les frais soient plus que couverts.

Embellir la perception sensorielle humaine

Là où l'IA excelle à améliorer l'interaction humaine, c'est notamment en augmentant les capacités des êtres

humains : en leur permettant d’utiliser leurs sens pour travailler avec des données augmentées, ou en augmentant ces sens pour que les gens puissent faire mieux. Les sections qui suivent traitent de ces deux manières de renforcer la perception sensorielle humaine et d’améliorer ainsi la communication.

Modifier le spectre des données

Souvent, pour rassembler divers types d’informations, nous employons des technologies qui filtrent ou modifient le spectre des données en fonction des couleurs, des sons ou des odeurs. Nous utilisons toujours nos capacités naturelles, mais une technologie donnée change l’input de manière à le rendre compatible avec nos capacités. Un des exemples les plus classiques de modification d’un spectre est l’astronomie : c’est en changeant et en filtrant la lumière que nous parvenons à observer des objets astronomiques invisibles à l’œil nu, par exemple des nébuleuses, et ainsi, à mieux comprendre l’univers. Or, changer et filtrer manuellement les couleurs, les sons et les odeurs peut demander beaucoup de temps, et les résultats peuvent être décevants quelles que soient les compétences exercées. C’est ici qu’intervient l’IA. Une IA peut essayer différentes combinaisons bien plus rapidement qu’un être humain et déterminer les combinaisons potentiellement utiles avec beaucoup plus de facilité, sachant qu’elle exécute la tâche de façon cohérente.



La technique la plus étonnante pour explorer notre monde est entièrement différente de ce que l’on imagine le plus souvent. Et si l’on pouvait sentir une couleur ou voir un son ? L’existence de la synesthésie (http://www.science20.com/news_releases/synaesthesia_smelling_a_sound_or_hearing_a_color), qui est l’utilisation d’un sens pour interpréter un input

associé à un autre sens, est bien documentée chez les humains. L'IA est utilisée pour étudier cet effet, comme l'explique l'article sur la page http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004959.
Cependant, cette technologie est surtout intéressante pour créer les conditions permettant aux gens de recourir à la synesthésie comme moyen différent de voir le monde (voir https://www.fastcompany.com/3024927/this-app-aids-yourdecision-making-by-mimicking-its-creators-synesthesia). Si vous voulez voir ce que le recours à la synesthésie peut vous apporter, intéressez-vous à l'application ChoiceMap à l'adresse https://choicemap.co/.

Augmenter les sens de l'être humain

Plutôt que d'utiliser une application externe pour modifier le spectre des données afin de produire des données utilisables par les humains, on peut augmenter les sens de l'être humain. Le principe de l'augmentation est qu'un appareil, externe ou implanté, permet à la personne de traiter directement et d'une nouvelle manière une information sensorielle. Beaucoup de gens assimilent ces nouvelles capacités à la création de cyborgs : voir https://www.theatlantic.com/technology/archive/2017/10/cyborg-future-artificialintelligence/543882/. L'idée n'est pas nouvelle : utiliser des outils pour rendre les êtres humains plus performants dans une série de tâches. Dans ce scénario, les gens bénéficient d'une augmentation sous deux formes : physique et intellectuelle.

L'augmentation physique des sens humains est déjà pratiquée de différentes manières et il est certain qu'elle

va se développer, à mesure que les gens deviendront plus disposés à recevoir divers types d'implants. Les lunettes de vision nocturne, par exemple, permettent aujourd'hui de voir la nuit, et les modèles haut de gamme offrent une vision en couleur contrôlée par un processeur spécialement conçu. Dans l'avenir, l'augmentation ou le remplacement des yeux pourraient permettre de voir n'importe quelle partie du spectre selon un processus contrôlé par la pensée. On ne verrait que la partie du spectre nécessaire pour accomplir une tâche donnée.

L'*Intelligence augmentée* suppose des mesures plus intrusives, mais c'est aussi la promesse de donner aux humains la possibilité d'exploiter des capacités bien plus grandes. Contrairement à l'IA, l'intelligence augmentée se caractérise par la présence d'un acteur humain au centre du traitement. L'être humain fournit la créativité et l'intention qui font pour le moment défaut à l'IA. Vous pouvez lire une étude des différences entre l'intelligence artificielle et l'intelligence augmentée à la page *https://www.usine-digitale.fr/article/de-l-intelligence-artificielle-a-l-intelligence-augmentee-la-nouvelle-ere-de-l-inte-raction-homme-machine.N584073*

PARTIE 3
Travailler avec des applications électroniques de l'IA

DANS CETTE PARTIE…

- Effectuer une analyse des données
- Étudier la relation entre l'IA et l'apprentissage machine
- Étudier la relation entre l'IA et l'apprentissage profond

Chapitre 9

Effectuer une analyse de données pour l'IA

DANS CE CHAPITRE

- » Comprendre comment fonctionne l'analyse de données
- » Exploiter efficacement l'analyse des données grâce à l'apprentissage machine
- » Déterminer ce que l'apprentissage machine permet d'obtenir
- » Découvrir les différents types d'algorithmes d'apprentissage machine

L'accumulation de données n'est pas un phénomène d'apparition récente : elle se pratique depuis des siècles. Que l'information apparaisse sous forme de texte ou sous forme numérique, les hommes ont toujours apprécié la façon dont elle décrit le monde environnant, et ils s'en servent pour faire progresser la civilisation. Les données ont une valeur propre. Elles permettent d'apprendre, de transmettre les informations essentielles à la génération suivante (ainsi il n'est pas nécessaire de réinventer la roue) et d'agir efficacement dans le monde.

On sait depuis quelque temps que les données contiennent davantage que des informations de surface. Quand elles sont sous une forme numérique adéquate, des techniques particulières mises au point par des mathématiciens et des statisticiens, appelées techniques d'analyse, permettent d'en extraire des informations

supplémentaires. En outre, à partir d'une simple analyse de données, il est possible d'extraire des informations utiles et de procéder à une analyse plus poussée à l'aide d'algorithmes d'apprentissage machine capables de prédire l'avenir, de classer les informations et de prendre des décisions.

L'analyse de données et l'apprentissage machine permettent de dépasser les limites précédentes en matière d'utilisation de données et de développer une IA plus performante. Ce chapitre est une introduction à l'analyse de données. Nous vous expliquons ici comment vous servir des données comme d'un outil d'apprentissage pour résoudre des problèmes difficiles d'IA, comme suggérer le bon produit à un consommateur, comprendre un langage parlé, traduire de l'anglais vers l'allemand, faire rouler une voiture sans conducteur, *etc*.

Définir l'analyse de données

Si notre époque est appelée l'ère de l'information, ce n'est pas simplement parce que les données sont devenues si abondantes, c'est aussi parce que la société a atteint une certaine maturité dans l'analyse et l'extraction de l'information. Google (avec sa filiale Alphabet), Amazon, Apple, Facebook et Microsoft, qui ont bâti leur activité sur les données, sont considérées comme les cinq plus grandes entreprises du monde. Ces compagnies ne se contentent pas de recueillir et de stocker les données produites par leurs processus numériques : elles savent les rendre aussi précieuses que le pétrole, grâce à une analyse de données précise et élaborée. Google, par exemple, enregistre des données sur l'Internet en général et à partir de son propre moteur de recherche, entre autres activités.

Dans la presse, les magazines et les conférences, on entend souvent affirmer que les données sont le pétrole d'aujourd'hui. Cette affirmation suppose que les données peuvent faire la fortune d'une entreprise, et qu'il faut pour cela des compétences et beaucoup de travail. Cette idée a été maintes fois reprise, mais c'est Clive Humby, un mathématicien britannique, qui, le premier, a comparé les données au pétrole, sur la base de son expérience des données des consommateurs dans le secteur de la vente au détail. Humby est connu pour avoir fait partie des fondateurs de Dunnhumby, une société britannique de marketing, et pour être l'éminence grise du programme de cartes de fidélité de Tesco. En 2006, Humby a aussi souligné que les données n'étaient pas simplement de l'argent qui tombe du ciel et qu'il fallait des efforts pour faire en sorte qu'elles soient utiles. De même qu'il n'est pas possible d'utiliser directement le pétrole non raffiné car il faut d'abord le transformer en essence, en matière plastique, en gaz ou autre produit chimique, les données doivent subir des transformations significatives pour pouvoir acquérir de la valeur.

Les transformations de données les plus élémentaires sont ce qu'on appelle l'analyse de données, et l'on peut les comparer aux transformations chimiques par lesquelles passe le pétrole dans une raffinerie avant de devenir un carburant ou une matière plastique utilisable. L'analyse de données permet de préparer le terrain à des processus d'analyse plus avancés, applicables aux données. Selon le contexte, l'analyse de données peut faire référence à diverses opérations sur les données, lesquelles opérations peuvent être spécifiques à certaines industries ou à certaines tâches. On peut classer toutes ces transformations selon quatre grandes familles, et cela donne une idée de ce qu'implique l'analyse de données :

» **Une transformation :** changer l'apparence des données. Le terme de transformation fait référence à

différents processus, le plus courant consistant à ordonner les données en lignes et en colonnes dans un *format matriciel* (on parle aussi de *fichier plat*). Concernant les produits achetés dans un supermarché, par exemple, on peut créer une ligne par client, et y représenter par des entrées numériques les produits que le client a achetés, en y faisant figurer les quantités ou les valeurs monétaires. Il peut s'agir aussi de transformations numériques comme le *redimensionnement*, par lequel on change la *moyenne* ou les valeurs extrêmes d'une série numérique afin de l'adapter à un algorithme.

» **Un nettoyage :** réparer ou éliminer les données imparfaites. Selon la façon dont les données ont été obtenues, différents problèmes peuvent se poser. Il peut y avoir des informations manquantes, des valeurs extrêmes, ou simplement des valeurs fausses. Ainsi, des données relatives à un supermarché peuvent présenter des erreurs, si certains prix sont incorrects. Il se peut que certaines données soient contradictoires ou qu'elles aient été créées pour fausser les conclusions, comme lorsque de faux avis sont publiés sur Internet pour donner une meilleure cote à un produit. La purification permet d'éliminer ces données illusoires et de rendre les conclusions plus fiables.

» **Un contrôle :** pour valider les données. L'analyse des données est essentiellement un travail humain, même si le logiciel joue un grand rôle. Les humains peuvent facilement reconnaître des tendances et repérer les éléments atypiques. Pour cette raison, l'analyse des données fournit de nombreuses statistiques et permet d'obtenir des représentations visuelles éclairantes, comme Health InfoScape, de Senseable Cities et General Electric

(http://senseable.mit.edu/healthinfoscape/), qui permet de saisir l'information rapidement. Cela permet de voir, par exemple, quel est le lien entre une maladie et une autre en se référant aux résultats du traitement de 72 millions d'enregistrements.

» **Une modélisation** : pour saisir la relation entre les éléments présents dans les données. Pour cela, il faut des outils statistiques, comme les corrélations, le test de Student, la régression linéaire et beaucoup d'autres, qui permettent de déterminer si une valeur est véritablement différente d'une autre ou simplement liée. Si l'on reprend l'exemple de l'analyse des dépenses au supermarché, on peut déterminer que souvent, les consommateurs qui achètent des couches-culottes achètent aussi de la bière. L'analyse statistique montre que ces deux produits sont souvent achetés ensemble (l'étude en question est devenue une vraie légende dans l'analyse des données : lisez cette courte histoire racontée dans un article du magazine Forbes à l'adresse https://www.forbes.com/global/1998/0406/0101102s1.html).

L'analyse de données n'a rien de magique. Il s'agit d'effectuer une transformation, un nettoyage, un contrôle et une modélisation des données au moyen d'un calcul matriciel (qui n'est rien d'autre qu'une longue série d'additions et de multiplications comme nous apprenons à les faire à l'école). Les moyens utilisés comportent également des outils statistiques comme la moyenne et la variance, qui décrivent la distribution des données, et des outils plus sophistiqués comme le calcul de corrélation et la régression linéaire, qui indiquent si des événements ou des phénomènes peuvent être liés les uns aux autres

(comme l'achat de couches-culottes et l'achat de bouteilles de bière) à partir des données observées.



Ce qui complique l'analyse de données à l'ère du « big data », ce sont ces énormes volumes de données pour le traitement desquels il est nécessaire d'utiliser des outils particuliers comme Hadoop (http://hadoop.apache.org/) ou Apache Spark (https://spark.apache.org/), deux outils logiciels conçus pour effectuer des opérations sur de grandes masses de données. Malgré ces outils perfectionnés, il reste nécessaire de préparer jusqu'à 80 % des données, ce qui représente un sacré travail. Une interview intéressante de Monica Rogati, qui est spécialiste de ce domaine et intervient comme conseil en matière d'IA auprès d'un certain nombre d'entreprises, traite de cette question plus en détail : https://www.nytimes.com/2014/08/18/technology/for-big-data-scientists-hurdle-to-insights-is-janitor-work.html.

Comprendre pourquoi l'analyse est importante

Dans le domaine de l'IA, l'analyse de données est essentielle. Il n'est même pas possible aujourd'hui d'envisager une IA sans que les données soient visualisées, nettoyées, transformées et modélisées avant que des algorithmes élaborés entrent en jeu pour en faire des informations plus utiles encore.

Au début, quand l'IA se résumait à des solutions purement algorithmiques et à des systèmes experts, les scientifiques et les spécialistes des données préparaient soigneusement les données pour alimenter le système. S'il fallait un algorithme pour trier les informations, par

exemple, un spécialiste des données les ordonnait sous forme de listes ou autres structures de données. On disposait ainsi de données appropriées se prêtant aux manipulations voulues. Les spécialistes des données recueillaient et organisaient les données de telle sorte que le contenu et la forme soient précisément adaptés au besoin spécifique en question. La manipulation de données connues pour les mettre sous une forme spécifique était contraignante, dans la mesure où elle demandait beaucoup de temps et d'énergie : les algorithmes recevaient moins d'informations qu'aujourd'hui.

Aujourd'hui, l'attention se porte moins sur la production de données que sur leur préparation au moyen de l'analyse de données. L'idée est que des sources variées produisent déjà des données en telles quantités que l'on peut trouver ce que l'on cherche sans être obligé de créer une série de données particulière pour la tâche en question. Imaginez, par exemple, qu'il vous faille une IA pour contrôler votre chatière, de telle sorte que seuls vos chiens et chats puissent entrer dans votre maison. Les algorithmes d'IA actuels apprennent à partir de données spécifiques à la tâche, ce qui signifie qu'ils traitent un grand nombre d'images montrant des chiens, des chats et d'autres animaux. De façon probable, cette énorme quantité d'images arrivera par Internet, en provenance de réseaux sociaux ou à partir de recherches d'images. Auparavant, il fallait pour cela que les algorithmes utilisent simplement quelques informations spécifiques concernant par exemple la forme, la taille et autres caractéristiques distinctives des animaux. Compte tenu de l'insuffisance de données, ils ne pouvaient exécuter qu'un nombre limité de tâches. En fait, il n'existe aucun exemple d'IA capable de commander une chatière en utilisant des algorithmes ou des systèmes experts classiques.

L'analyse de données vient à la rescousse des algorithmes actuels en produisant des informations sur les images récupérées sur l'Internet. L'utilisation de l'analyse de données permet à l'IA de découvrir la taille des images, leur variété, le nombre de couleurs, les mots utilisés dans les titres des images, *etc*. Elle fait partie du contrôle des données, et, en l'occurrence, de ce qui est nécessaire pour les nettoyer et les transformer. L'analyse de données peut vous permettre, par exemple, de trouver la photo d'un animal étiqueté par erreur comme un chat (il ne s'agit pas d'induire en erreur votre IA) et de transformer les images de manière à utiliser un format de couleurs unique (par exemple, des nuances de gris) et une taille unique.

Reconsidérer l'utilité des données

Avec l'explosion de la quantité de données disponibles sur les appareils numériques (voir Chapitre 2), les données trouvent une nouvelle valeur et une nouvelle utilité au-delà de l'objectif initial d'instruction (enseignement) et de transmission de la connaissance (transfert de données). L'abondance des données qui alimentent l'analyse de données acquiert de nouvelles fonctions en dehors des fonctions informatives :

» Les données décrivent mieux le monde en présentant une grande variété de faits, et elles le décrivent plus en détail en révélant des nuances pour chaque fait. Elles sont devenues si abondantes qu'elles couvrent tous les aspects de la réalité. Elles peuvent servir à découvrir comment des choses ou des faits apparemment sans lien sont en réalité liés.

» Les données montrent comment les faits sont associés aux événements. On peut en tirer des règles générales et des enseignements sur la

> manière dont le monde va changer ou se transformer, compte tenu de certaines prémisses. Quand les gens agissent d'une certaine façon, les données offrent aussi une certaine capacité prédictive.

Par certains aspects, les données nous apportent de nouveaux superpouvoirs. Chris Anderson, l'ancien rédacteur en chef de *Wired*, explique comment de grandes quantités de données favorisent les découvertes scientifiques indépendamment de la méthode scientifique (lire l'article à l'adresse https://www.wired.com/2008/06/pb-theory/). Cet auteur s'appuie sur l'exemple de la réussite de Google dans les domaines de la publicité et de la traduction, une réussite obtenue non pas en utilisant des modèles ou des théories spécifiques, mais en appliquant des algorithmes pour tirer des enseignements à partir des données.

Comme dans la publicité, des données scientifiques (concernant par exemple la physique ou la biologie) peuvent favoriser une innovation permettant aux scientifiques d'aborder les problèmes autrement qu'avec des hypothèses, en étudiant les variations observées dans de grands volumes de données et en utilisant des algorithmes de découverte. Galilée avait utilisé la méthode scientifique pour créer les fondements de la physique et de l'astronomie modernes (voir https://www.biography.com/people/galileo-9305220). Les premières avancées reposent la plupart du temps sur des observations et sur des expérimentations contrôlées qui expliquent comment et pourquoi les phénomènes se produisent. La possibilité d'innover en utilisant les seules données constitue une avancée majeure dans notre manière de comprendre le monde.

Dans le passé, les scientifiques effectuaient d'innombrables observations et procédaient à une

multitude de déductions pour décrire la physique de l'univers. Ce processus manuel permettait de trouver les lois sous-jacentes du monde dans lequel nous vivons. L'analyse de données, en associant des observations exprimées sous forme d'intrants et d'extrants, nous permet de déterminer la façon dont les choses fonctionnent et de définir, grâce à l'apprentissage machine, des règles ou des lois approximatives sans devoir recourir à l'utilisation d'observations et de déductions manuelles. Le processus est maintenant plus rapide et plus automatisé.

Définir l'apprentissage machine

Dans l'analyse de données, rien n'est aussi grandiose que l'apprentissage machine. Pour pouvoir l'exploiter utilement, il faut d'abord que l'analyse de données produise un ensemble d'informations qui soit correct. Cependant, seul l'apprentissage machine peut associer une série d'extrants à une série d'intrants et déterminer de façon probante les règles de fonctionnement qui aboutissent à ces extrants. L'analyse de données porte surtout sur la compréhension et la manipulation des données, qui doivent devenir plus exploitables et fournir un éclairage sur le monde, tandis que l'apprentissage machine consiste exclusivement à tirer des informations à partir des données et à élaborer une représentation interne du monde qui soit exploitable à des fins pratiques. L'apprentissage machine permet de prédire l'avenir, d'effectuer un classement pertinent ou de prendre la meilleure décision rationnelle dans un contexte donné.

S'APERCEVOIR QUE LA QUALITÉ DE L'IA DÉPEND DES DONNÉES

Les données ne font pas qu'alimenter l'IA, elles sont aussi ce qui la rend possible. On pourrait dire que l'IA est le résultat d'algorithmes élaborés à partir d'une base mathématique complexe. Des activités comme la vision et la compréhension du langage supposent des algorithmes qu'il n'est pas facile d'expliquer en termes simples et qui mettent en jeu des millions de calculs (le matériel informatique a aussi son rôle à jouer).

Cependant, l'IA, ce ne sont pas seulement des algorithmes. Alexander Wissner-Gross, un scientifique américain qui est chercheur, entrepreneur et qui enseigne à l'Institute for Applied Computation Science à Harvard, donne son point de vue dans une récente interview pour Edge (https://www.edge.org/response-detail/26587). Il se demande pourquoi il a fallu tant de temps pour que l'IA commence vraiment à se développer, et conclut que c'était peut-être un problème de qualité et de disponibilité des données plutôt que de capacités algorithmiques. Wissner-Gross étudie l'ordonnancement des avancées en matière d'IA au cours de ces dernières années et montre comment les données et les algorithmes contribuent au succès de chaque percée, en soulignant que chacune de ces percées était récente au moment où le tournant a été atteint. Il fait valoir que les données sont relativement nouvelles et toujours actualisées, tandis que les algorithmes ne sont pas de nouvelles découvertes, qu'ils sont plutôt l'amélioration d'une technologie plus ancienne.

Wissner-Gross conclut qu'en moyenne, l'algorithme est de 15 ans plus ancien que les données. Il fait valoir que les données favorisent les progrès, et invite le lecteur à se demander ce qui pourrait se passer s'il était possible d'alimenter les algorithmes dont on peut disposer actuellement

avec des données qui seraient meilleures en termes de qualité et de quantité.



L'idée qui sous-tend l'apprentissage machine est que l'on peut représenter la réalité en utilisant une fonction mathématique que l'algorithme ne connaît pas d'avance, mais qu'il peut déduire de l'examen d'une série de données. On peut exprimer la réalité dans toute sa redoutable complexité sous forme de fonctions mathématiques inconnues que les algorithmes d'apprentissage machine trouvent et rendent disponibles. Ce concept est le plus essentiel pour tous les algorithmes d'apprentissage machine.

Dans l'apprentissage machine, l'apprentissage est purement mathématique et il se termine par l'association de certains intrants à certains extrants. Il n'a rien à voir avec la compréhension de ce que l'algorithme a appris (l'analyse de données n'améliore la compréhension que dans une certaine mesure), aussi le processus d'apprentissage est-il souvent considéré comme une formation, sachant que l'algorithme a appris à associer la réponse correcte (l'extrant) à chaque question soumise (l'intrant).

Bien qu'il ne s'agisse pas véritablement de compréhension mais d'un simple processus mathématique, l'apprentissage machine se révèle utile dans bien des tâches. Il donne à l'IA le pouvoir de faire ce qui est le plus rationnel dans un contexte donné, où l'apprentissage se fait en utilisant les bonnes données. Les sections qui suivent permettent de comprendre plus précisément comment fonctionne l'apprentissage machine, quels avantages on peut espérer en tirer, et quelles sont les limites de l'utilisation de l'apprentissage machine dans une application.

Comprendre comment fonctionne l'apprentissage machine

Beaucoup de gens ont l'habitude de penser qu'une application se résume à une fonction qui accepte des données en entrée et produit un résultat. Un programmeur pourra créer une fonction qu'il appellera `Add ()`, qui acceptera en entrée deux valeurs, par exemple 1 et 2. Le résultat de `Add ()` sera 3. Le résultat du processus est une valeur. Dans le passé, pour écrire un programme, il fallait comprendre la fonction utilisée pour manipuler les données entrées de manière à obtenir un certain résultat. Avec l'apprentissage machine, le processus est inversé. Vous connaissez vos deux entrées, par exemple 1 et 2, et vous savez aussi que le résultat désiré est 3. Ce que vous ne savez pas, c'est quelle fonction appliquer pour obtenir ce résultat. L'apprentissage consiste à fournir un algorithme avec toutes sortes d'exemples de données entrantes et de résultats attendus à partir de ces entrées. L'apprenant utilise ces données pour créer une fonction. En d'autres termes, l'apprentissage est le processus par lequel l'algorithme de l'apprenant élabore une fonction flexible adaptée aux données. Le résultat est généralement la probabilité d'une certaine classe ou d'une certaine valeur numérique.

Pour avoir une idée plus nette de ce processus, imaginons un enfant qui apprend à distinguer les arbres des autres objets. Avant qu'il soit capable de le faire seul, l'enseignant lui montre un certain nombre d'images d'arbres sur lesquelles figurent tous les éléments distinctifs des arbres : la matière (le bois), les parties de l'arbre (le tronc, les branches, les feuilles ou les aiguilles, les racines), et le lieu (les arbres sont plantés dans le sol). L'enfant se fait ainsi une idée de l'aspect d'un arbre grâce

au contraste entre ces caractéristiques et les images d'objets différents comme un meuble, qui est en bois mais ne présente pas les autres caractéristiques des arbres.

Le classifieur de l'apprentissage machine fonctionne de la même manière. Il développe ses capacités cognitives en créant une formulation mathématique qui intègre tous les éléments de manière à obtenir une fonction capable de distinguer une catégorie d'une autre. Supposons qu'il existe une formulation mathématique, aussi appelée fonction cible, exprimant les caractéristiques d'un arbre. Le classifieur peut alors chercher une représentation sous forme de réplique ou d'approximation (une fonction différente dont le fonctionnement est semblable). La capacité d'exprimer cette formulation mathématique, c'est la capacité de représentation du classifieur.

D'un point de vue mathématique, on peut exprimer le processus de représentation dans l'apprentissage machine en modélisant un terme équivalent. Il y a modélisation quand on découvre la construction d'une fonction en observant ses résultats. Dans l'apprentissage machine, le processus de modélisation est similaire à celui par lequel un enfant assimile l'idée d'un objet. Il comprend les règles abstraites tirées des données du monde réel, si bien qu'il reconnaît un arbre dès qu'il en voit un.

Si une telle représentation (des règles abstraites tirées des données du monde réel) est possible, c'est parce que l'algorithme d'apprentissage comporte un certain nombre de paramètres (constitués de vecteurs et de matrices), qui correspondent à sa mémoire et gèrent des idées appropriées à son activité de modélisation consistant à relier des propriétés à des classes de réponse. Les dimensions et le type de ces paramètres internes délimitent le type de fonctions cibles qu'un algorithme peut apprendre. Dans cet algorithme, un moteur d'optimisation change les paramètres au cours de l'apprentissage pour représenter la fonction cachée de la cible.

Au cours de l'optimisation, l'algorithme cherche les variantes possibles de ses combinaisons de paramètres afin d'en trouver une qui permette une modélisation correcte entre les propriétés et les classes pendant l'apprentissage. Ce processus évalue un certain nombre de fonctions cibles candidates parmi celles que l'algorithme d'apprentissage peut prévoir. L'ensemble des fonctions potentielles que l'algorithme est capable de découvrir s'appelle l'*espace d'hypothèses*. Le classifieur résultat, avec sa série de paramètres, peut être considéré comme une hypothèse, une façon de dire que l'algorithme a défini des paramètres pour reproduire la fonction cible et qu'il est maintenant prêt pour définir des classifications correctes (un fait démontré par la suite).

L'espace d'hypothèses doit contenir toutes les variantes des paramètres de tous les algorithmes d'apprentissage machine que l'on veut modéliser en association avec une fonction inconnue quand on désire résoudre un problème de classification. Des algorithmes différents peuvent avoir des espaces d'hypothèses différents. Ce qui importe vraiment, c'est que l'espace d'hypothèses contienne la fonction cible (ou son approximation, qui est une fonction différente mais similaire, car tout ce dont on a besoin en fin de compte, c'est d'obtenir quelque chose qui fonctionne).

On peut concevoir cette phase comme le moment où l'enfant expérimente diverses idées créatives en rassemblant ses connaissances et ses expériences (une analogie avec les propriétés données) pour créer la visualisation d'un arbre. Naturellement, ses parents interviennent dans cette phase et ce sont eux qui assurent la présence des éléments environnementaux adéquats. Dans l'apprentissage machine, il faut que quelqu'un fournisse les bons algorithmes d'apprentissage ainsi que des paramètres qui ne peuvent pas être appris (on les appelle les hyperparamètres), choisisse une série

d'exemples à partir desquels se fera l'apprentissage, et établisse les propriétés liées à ces exemples. De même qu'un enfant ne peut pas toujours apprendre à distinguer le bien et le mal quand il est livré à lui-même, les algorithmes d'apprentissage machine ont besoin des êtres humains pour pouvoir apprendre.

Percevoir l'intérêt de l'apprentissage machine

L'IA et l'apprentissage machine sont aujourd'hui utilisés dans tout un ensemble d'applications. Le seul problème est que la technologie étant si performante, on ne sait même pas qu'elle existe. Vous seriez peut-être surpris d'apprendre que ces deux technologies sont déjà présentes dans des appareils que vous utilisez chez vous. Elles sont présentes à coup sûr dans votre voiture, ainsi que sur votre lieu de travail. En fait, elles sont extrêmement répandues, mais on ne les voit pas, même quand elles sont spectaculaires. Voici quelques-unes parmi d'autres des applications les moins invisibles :

» **La détection des fraudes :** supposons que votre compagnie de cartes de crédit vous appelle pour vérifier que c'est bien vous qui avez effectué un achat particulier. Ce n'est pas la compagnie qui a du flair, c'est simplement que L'IA a détecté une dépense inhabituelle et a déclenché une alerte.

» **La planification des ressources :** nombreuses sont les organisations qui doivent planifier leurs ressources pour pouvoir fonctionner de façon efficiente. Un hôpital, par exemple, peut avoir besoin de déterminer où installer un patient en tenant compte des besoins de celui-ci, de la disponibilité des spécialistes, et du temps pendant lequel les

médecins estiment que le patient va devoir rester hospitalisé.

» **L'analyse complexe :** les humains ont souvent besoin de s'appuyer sur une analyse complexe, parce qu'il y a trop de facteurs à prendre en compte. Par exemple, les mêmes symptômes peuvent indiquer plusieurs troubles de la santé. Les médecins doivent parfois recourir à un processus complexe afin d'être en mesure de formuler le bon diagnostic à temps pour pouvoir sauver la vie du patient.

» **L'automation :** l'ajout d'une IA peut être profitable à toute forme d'automation pour faire face à des changements ou événements imprévus. Un problème qui se pose aujourd'hui avec certains types d'automation est qu'un événement inattendu, par exemple un objet qui n'est pas à sa place, peut entraîner une panne. En associant une IA à l'automation, on peut lui permettre de gérer l'imprévu et d'assurer la poursuite du processus comme s'il ne s'était rien passé.

» **Le service à la clientèle :** aujourd'hui, quand vous téléphonez au service clients d'une société, il se peut qu'il n'y ait plus aucun être humain à l'autre extrémité de la ligne. L'automation est assez élaborée pour qu'un système automatisé, en suivant des scripts et en utilisant diverses ressources, puisse répondre à la plupart de vos questions. Si les inflexions de voix (également grâce à l'IA) sont bien conçues, vous ne saurez même pas que c'est à un ordinateur que vous parlez.

» **Les systèmes de sécurité :** aujourd'hui, diverses sortes de machines sont équipées de systèmes de sécurité faisant intervenir l'IA. Il en est ainsi, par exemple, du système de freinage automatique qui

immobilise un véhicule en cas de nécessité, en fonction de données comme le sens du dérapage.

» **L'efficacité des machines :** l'IA permet de contrôler une machine de manière à en maximiser l'efficience. Elle contrôle l'utilisation des ressources afin d'éviter, par exemple, un dépassement de la vitesse autorisée. La moindre once d'énergie est utilisée dans la mesure du nécessaire pour assurer les services voulus.

Cette liste est à peine un léger aperçu de toutes les possibilités existantes. L'IA trouve bien d'autres utilisations encore. Cependant, il est utile également d'envisager l'utilisation de l'apprentissage machine en dehors de ce qui est souvent considéré comme le domaine habituel de l'IA. Voici quelques exemples d'utilisation de l'apprentissage machine que vous n'auriez peut-être pas eu l'idée d'associer à l'IA :

» **Le contrôle d'accès :** bien souvent, le contrôle d'accès est une question à laquelle la réponse est oui ou non. Une carte à puce donne à un salarié l'accès à une ressource, de la même manière qu'une clé ouvrant une serrure. Certains systèmes d'accès présentent la possibilité de régler les horaires et les dates d'accès autorisé, mais ce type de contrôle assez grossier ne répond pas vraiment à tous les besoins. Grâce à l'apprentissage machine, vous pouvez déterminer si une personne doit avoir ou non accès à une ressource, en fonction de son rôle et du besoin. Ainsi, un employé pourra avoir accès à une salle de formation si et seulement si la formation est en lien avec sa fonction dans l'entreprise.

» **La protection des animaux :** l'océan peut sembler assez vaste pour que les animaux marins et les bateaux cohabitent sans problème. Malheureusement, il arrive souvent qu'un mammifère

marin soit heurté par un bateau. Un algorithme d'apprentissage machine peut permettre aux bateaux d'éviter ces collisions en apprenant les sons émis par les animaux et ceux émis par les bateaux, ainsi que les caractéristiques des uns et des autres.

» **La prédiction des temps d'attente :** la plupart des gens n'aiment pas attendre quand ils n'ont aucune idée du temps que durera l'attente. Avec l'apprentissage machine, une application peut déterminer les temps d'attente en fonction du niveau des effectifs, de la charge de travail du personnel, de la complexité des problèmes que le personnel essaie de résoudre, de la disponibilité des ressources, *etc*.

C'est monotone mais c'est utile

Même si le cinéma invite à imaginer que l'IA ne peut que créer une énorme onde de choc, et même s'il vous arrive d'être témoin d'applications remarquables de l'IA dans la réalité, les utilisations les plus fréquentes de l'IA sont triviales, et même ennuyeuses. Ainsi, un article récent explique comment Verizon utilise le langage R dans le cadre de l'apprentissage machine pour l'analyse des données relatives aux défauts de sécurité et le suivi annuel automatisé de la sécurité (https://www.computerworld.com/article/3001832/dataanalytics/how-verizon-analyzes-security-breach-data-with-r.html). Cette analyse est un travail rébarbatif comparé à d'autres activités dans lesquelles l'IA joue un rôle, mais Verizon réalise des économies substantielles en effectuant cette analyse à l'aide du langage R, et les résultats n'en sont que meilleurs.

Spécifier les limites de l’apprentissage machine

L’apprentissage machine recourt à des algorithmes pour analyser d’énormes ensembles de données. Actuellement, il ne peut pas produire le genre d’IA que montrent les films. Même les meilleurs algorithmes ne peuvent pas penser, sentir, manifester une forme ou une autre de conscience de soi, ni exercer un quelconque libre arbitre. Ce que peut faire l’apprentissage machine, c’est effectuer des analyses prédictives bien plus rapidement que n’importe quel être humain. En conséquence, l’apprentissage machine nous permet de travailler plus efficacement. Le stade actuel de l’IA, c’est donc celui de la réalisation d’analyses, mais les humains doivent encore prendre en compte les implications de ces analyses et prendre des décisions en ne négligeant pas les aspects moraux et éthiques. En principe, l’apprentissage machine ne constitue que la partie apprentissage de l’IA, et cette partie est loin d’être susceptible de donner naissance au genre d’IA qu’on voit dans les films.

Le principal point de confusion entre apprentissage et intelligence réside dans le fait de supposer simplement que lorsqu’une machine exécute une tâche de mieux en mieux (lorsqu’elle apprend), elle a aussi une conscience (une intelligence). Or, concernant l’apprentissage machine, rien ne confirme ce point de vue. Il se produit le même phénomène quand les gens supposent qu’un ordinateur peut leur compliquer la vie exprès. L’ordinateur ne peut pas gérer des émotions, il ne peut qu’agir en fonction de ce qui lui est fourni en entrée et des instructions du programme exécuté pour traiter cette entrée. Il existera une vraie IA le jour où les ordinateurs finiront par être capables d’émuler cette combinaison subtile que la nature a réalisée :

- **Génétique :** apprentissage lent d'une génération à la suivante.
- **Apprentissage :** apprentissage rapide à partir de sources organisées.
- **Exploration :** apprentissage spontané grâce à des supports et moyens de communication et à des interactions avec les autres.

En dehors du fait que l'apprentissage machine est constitué de fonctions mathématiques optimisées à certaines fins, d'autres faiblesses en font apparaître les limites, en particulier les trois suivantes :

- **Représentation** : représenter des problèmes à l'aide de fonctions mathématiques n'est pas facile, surtout lorsqu'il s'agit de problèmes complexes comme l'imitation du cerveau humain. Pour le moment, l'apprentissage machine ne peut résoudre que des problèmes simples et spécifiques, consistant à répondre à des questions simples comme « Qu'est-ce que c'est ? », « Combien ça coûte ? » ou « Qu'y a-t-il ensuite ? ».
- **Surapprentissage :** les algorithmes d'apprentissage machine semblent apprendre ce que vous vous souciez de leur faire apprendre, mais ce n'est pas la réalité. Leurs fonctions internes mémorisent les données sans apprendre de ces données. Il y a surapprentissage quand votre algorithme apprend trop de vos données, au point de créer des fonctions et des règles qui n'existent pas dans la réalité.
- **Absence de généralisation effective en raison de données limitées :** l'algorithme apprend ce qu'on lui enseigne. Si vous l'alimentez en données fausses ou anormales, il se comportera de façon anormale.

En termes de représentation, un seul algorithme d'apprentissage peut apprendre un certain nombre de choses différentes, mais un algorithme n'est pas adapté à toutes les tâches. Certains algorithmes ont un caractère assez général pour pouvoir jouer aux échecs, reconnaître des visages sur Facebook ou diagnostiquer un cancer chez un patient. Un algorithme réduit les entrées de données et les résultats attendus à une fonction dans tous les cas, mais la fonction est spécifique au type de tâche qu'il s'agit de faire exécuter par l'algorithme.

Le secret de l'apprentissage machine est la généralisation. Cependant, la généralisation s'accompagne de problèmes de surapprentissage et de données biaisées. L'objectif est de généraliser la fonction de sortie de telle manière qu'elle traite les données au-delà des exemples contenus dans l'apprentissage. Prenons l'exemple d'un filtre antispam. Supposons que votre dictionnaire contienne 100000 mots (un petit dictionnaire). Un jeu de données limité pour l'apprentissage, de 4000 à 5000 combinaisons de mots, doit donner lieu à une fonction généralisée capable de détecter le spam dans les 2 puissance 100000 combinaisons qu'elle passera en revue en traitant les données réelles. Dans de telles conditions, on aura l'impression que l'algorithme apprend les règles du langage, mais ce ne sera pas la réalité. L'algorithme pourra réagir correctement à des situations similaires à celles qui ont fait partie de l'apprentissage, mais il ne saura pas quoi faire dans des situations entièrement nouvelles. Ou bien, il pourra présenter des biais de façon inattendue en raison du type de données qu'il recevait habituellement au cours de l'apprentissage.

Microsoft, par exemple, a appris à Tay, une IA, à dialoguer avec les utilisateurs sur Twitter et à s'instruire de leurs réponses. Malheureusement, les interactions ont « dérapé », car certains utilisateurs ont alimenté Tay avec des propos haineux, ce qui a suscité des inquiétudes

concernant le bien-fondé d'une IA reposant sur une technologie d'apprentissage machine (pour plus de détails sur cette anecdote, consultez la page https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbotracist).

Le problème est que l'algorithme d'apprentissage machine était alimenté par des données inopportunes et non filtrées (Microsoft n'a pas recouru à une analyse de données appropriée pour nettoyer et contrôler les intrants comme il aurait fallu le faire). Il en est résulté une certaine forme de surapprentissage, avec pour conséquence la sélection d'une mauvaise série incitant à la haine, plutôt que de représenter le monde d'une façon plus générale. D'autres IA destinées à dialoguer avec des utilisateurs, comme Mitsuku (http://www.mitsuku.com/), qui a été primée, ne sont pas exposées aux mêmes risques que Tay car leur apprentissage est strictement contrôlé et supervisé au moyen d'une analyse de données et d'une évaluation effectuée par des personnes humaines.

Savoir comment apprendre à partir des données

Dans l'apprentissage machine, tout tourne autour des algorithmes. Un algorithme est une procédure ou une formule utilisée pour résoudre un problème. Du domaine concerné dépend le type d'algorithme nécessaire, mais l'objectif de base est toujours le même : résoudre un problème, par exemple conduire une voiture, ou jouer aux dominos. Dans le premier cas, les problèmes sont complexes et nombreux, mais il s'agit en fin de compte de transporter un passager d'un endroit à un autre sans provoquer de collision. De même, quand on joue aux dominos, il y a un objectif essentiel : gagner la partie.

Différentes formes d'apprentissage sont possibles, selon l'algorithme et ses objectifs. On peut distinguer trois grandes catégories d'algorithmes d'apprentissage machine, en fonction de leur objet :

- l'apprentissage supervisé ;
- l'apprentissage non supervisé ;
- l'apprentissage par renforcement.

Les sections qui suivent expliquent plus en détail quelles différentes sortes d'algorithmes sont exploités par l'apprentissage machine.

L'apprentissage supervisé

On parle d'*apprentissage supervis*é quand un algorithme apprend à partir de données entrées à titre d'exemple et de réponses cibles associées qui peuvent être des valeurs numériques, ou bien des catégories ou des étiquettes, en vue de prédire par la suite la réponse correcte quand on l'alimente avec de nouveaux exemples. Cette approche est similaire à l'apprentissage humain sous la supervision d'un enseignant. L'enseignant donne à l'élève de bons exemples à mémoriser, et l'élève en déduit des règles générales.

Il convient de bien faire la distinction entre les *problèmes de régression*, dont la cible est une valeur numérique, et les *problèmes de classification*, dont la cible est une variable qualitative, par exemple une classe ou une étiquette. Une régression peut servir à déterminer les prix moyens des maisons dans la région de Boston, tandis qu'une classification consistera, par exemple, à distinguer plusieurs sortes de fleurs d'iris d'après les dimensions de leurs sépales et de leurs pétales. Voici quelques exemples d'apprentissage supervisé ayant d'importantes applications en IA, décrits en termes de données

entrantes, de données sortantes et d'application concrète à la résolution de problèmes :

Données entrantes (X)	Données sortantes (y)	Application
Historique des achats effectués par les consommateurs	Une liste de produits que les consommateurs n'ont jamais achetés	Système de recommandation
Images	Une liste de cases étiquetées avec le nom d'un objet	Détection et reconnaissance d'images
Texte en anglais sous forme de questions	Texte en anglais sous forme de réponses	Chatbot, application logicielle capable de converser
Texte en anglais	Texte en allemand	Traduction en langage machine
Audio	Transcription de texte	Reconnaissance vocale
Image, données des capteurs	Direction, braquage ou accélération	Planification comportementale pour une conduite autonome

L'apprentissage non supervisé

Il y a *apprentissage non supervisé* quand un algorithme apprend à partir d'exemples simples sans réponse associée et détermine lui-même les structures de données. Ce type d'algorithme restructure généralement les données en leur donnant une autre forme, celle de nouvelles propriétés qui peuvent représenter une classe ou une nouvelle série de valeurs non corrélées. Les

données résultantes sont très utiles pour éclairer l'utilisateur humain sur la signification des données initiales et produire des intrants exploitables par les algorithmes d'apprentissage machine supervisé.

L'apprentissage non supervisé ressemble aux méthodes utilisées par les humains pour déterminer que certains objets ou événements font partie d'une même classe et pour observer le degré de similarité entre eux. Certains systèmes de recommandation que l'on trouve sur Internet sous forme d'automatisation du marketing sont basés sur ce type d'apprentissage. L'algorithme d'automatisation du marketing élabore ses suggestions à partir de ce que vous avez acheté dans le passé. Les recommandations se fondent sur une estimation du groupe de clients auquel vous ressemblez le plus et sur une inférence consécutive de vos préférences probables en fonction de ce groupe.

L'apprentissage par renforcement

Il y a *apprentissage par renforcement* quand on soumet à l'algorithme des exemples sans étiquettes, comme dans l'apprentissage non supervisé. Cependant, un exemple peut s'accompagner d'une appréciation positive ou négative selon la solution que l'algorithme propose.

L'apprentissage par renforcement est lié aux applications pour lesquelles l'algorithme doit prendre des décisions (le produit est donc prescriptif, et non pas simplement descriptif comme dans l'apprentissage non supervisé), et les décisions ont des conséquences. Dans le monde des humains, il n'y a pas de différence avec l'apprentissage par tâtonnements. Les erreurs permettent d'apprendre, compte tenu de la sanction qui les accompagne (en termes de coût, de perte de temps, de regret, de peine,

etc.). On apprend ainsi qu'une certaine ligne de conduite a moins de chances de réussir que les autres. Un exemple intéressant d'apprentissage par renforcement est celui des ordinateurs qui apprennent eux-mêmes à jouer à des jeux vidéo.

Dans ce cas, une application fournit à l'algorithme des exemples de situations spécifiques. Par exemple, le joueur s'est perdu dans un labyrinthe en évitant un ennemi. L'application renseigne l'algorithme sur le résultat des actions qu'il peut effectuer, et l'algorithme apprend en essayant d'éviter ce qu'il découvre être dangereux et de faire en sorte de survivre. Pour voir comment DeepMind a créé un programme d'apprentissage par renforcement qui joue à des jeux vidéo d'un ancien Atari, consultez la page https://www.youtube.com/watch?v=V1eYniJ0Rnk. En regardant la vidéo, vous remarquerez que le programme est maladroit au début mais s'améliore peu à peu grâce à l'apprentissage, jusqu'à devenir champion. Les points forts et les points faibles du processus sont décrits par Raia Hadsell, une scientifique chevronnée de l'équipe de Deep Learning chez DeepMind, dans une vidéo instructive de TEDx Talks : https://www.youtube.com/watch?v=mqma6GpM7vM.

Chapitre 10

Utiliser l'apprentissage machine dans l'IA

DANS CE CHAPITRE

» **Utiliser les outils de différentes tribus dans le cadre d'un apprentissage à partir des données**

» **Découvrir comment les probabilités sont avantageuses pour l'IA**

» **Prédire en utilisant le modèle bayésien naïf et les réseaux bayésiens**

» **Segmenter les données en branches et en feuilles dans des arbres de décision**

L'apprentissage est depuis le début une partie importante de l'IA, car l'IA peut imiter un niveau d'intelligence comparable à celui de l'homme. Il a fallu beaucoup de temps et le recours à une variété d'approches pour parvenir à une forme d'imitation qui imite effectivement l'apprentissage. Aujourd'hui, l'apprentissage machine peut se vanter d'avoir atteint un niveau quasi humain pour des tâches spécifiques comme la classification des images ou le traitement du son, et il s'efforce d'atteindre un niveau similaire dans un certain nombre d'autres tâches.

L'apprentissage machine n'est pas complètement automatisé. On ne peut pas demander à l'ordinateur de lire un livre et s'attendre à ce qu'il comprenne quelque chose. L'automation signifie que l'ordinateur est capable

d’apprendre à se programmer lui-même pour exécuter des tâches, au lieu d’attendre que des humains le programment. Actuellement, l’automation nécessite de vastes volumes de données sélectionnées par l’humain, mais aussi une analyse de données et un apprentissage (là encore, sous la supervision des humains). Tout se passe comme quand on prend un enfant par la main pour le guider dans ses premiers pas. Par ailleurs, l’apprentissage machine présente d’autres limites, qui sont liées à la manière dont il apprend à partir des données.

Chaque famille d’algorithmes se caractérise par des méthodes spécifiques d’exécution des tâches, qui sont décrites dans ce chapitre. L’objectif est de comprendre comment l’IA prend des décisions et élabore des prédictions. Dans ce chapitre, comme on découvre l’homme derrière le rideau dans *Le Magicien d’Oz*, on découvre la machinerie et l’opérateur derrière l’IA. Toujours est-il que vous allez profiter de la sensation formidable que procure la découverte des merveilleuses réalisations rendues possibles par l’apprentissage machine.

Emprunter différents chemins vers l’apprentissage

Tout comme il existe chez les humains différentes manières d’apprendre à connaître le monde, les scientifiques qui ont étudié le problème de l’apprentissage dans le contexte de l’IA ont suivi différentes voies. Chacun a misé sur une recette particulière pour imiter l’intelligence. Jusqu’à présent, aucun modèle ne s’est révélé supérieur à un autre. Le théorème selon lequel rien n’est gratuit et chaque avantage se paye est pleinement vérifié. Chacune de ces tentatives s’est révélée efficace pour résoudre des problèmes particuliers. Les algorithmes

étant équivalents dans l'abstrait (voir l'encadré « On n'a rien sans rien »), aucun algorithme n'est supérieur à un autre, sauf preuve du contraire dans le cas d'un problème pratique spécifique. Les sections qui suivent présentent des informations complémentaires sur ce principe d'utilisation de différentes méthodes pour apprendre.

Découvrir cinq approches principales de l'apprentissage de l'IA

Un algorithme est un peu comme une boîte, dans la mesure où il stocke une méthode pour résoudre un type de problème particulier. Les algorithmes traitent les données en leur faisant franchir une série d'états qui, sans être nécessairement déterministes, sont bien définis. L'objectif est de créer un résultat qui résolve un problème. Dans certains cas, l'algorithme reçoit des intrants qui permettent de définir le résultat, mais l'accent est toujours mis sur le résultat.

Les algorithmes doivent exprimer les transitions entre les états dans un langage formel bien défini que l'ordinateur puisse comprendre. Pour traiter les données et résoudre le problème, l'algorithme définit, affine et exécute une fonction. La fonction est toujours spécifique au type de problème que l'algorithme doit traiter.

Comme l'indique la section « Éviter l'emphase concernant l'IA » du Chapitre 1, chacune des cinq tribus emploie une technique et une stratégie différentes pour résoudre les problèmes, et cela donne des algorithmes uniques. En combinant ces algorithmes, on devra, au bout du compte, obtenir l'algorithme maître qui sera capable de résoudre n'importe quel problème. Les sections qui suivent

présentent un aperçu des cinq principales techniques algorithmiques.

ON N'A RIEN SANS RIEN

Chez les mathématiciens, un théorème bien connu est celui de David Wolpert et William Macready, intitulé « No Free Lunch », selon lequel deux algorithmes d'optimisation sont toujours équivalents quand on fait la moyenne de leur performance sur l'ensemble des problèmes possibles. Fondamentalement, quel que soit l'algorithme d'optimisation utilisé, il n'offrira aucun avantage à être utilisé pour résoudre tous les problèmes possibles. Pour obtenir un avantage, il faut l'appliquer aux problèmes pour lesquels il excelle. L'article de Yo-Chi Ho et David L. Pepyne à l'adresse https://www.researchgate.net/publication/3934675_Simple_explanation_of_the_no_free_lunch_theorem_of_optimiza-tion donne une explication compréhensible mais rigoureuse de ce théorème. Pour plus de détails à propos des théorèmes de ce type, il est aussi judicieux de lire l'étude suivante : http://www.no-free-lunch.org/. L'apprentissage machine repose sur deux de ces théorèmes.

Le raisonnement symbolique

Une des premières tribus, celle des symbolistes, considérait que la connaissance peut être obtenue en effectuant des opérations sur des symboles (des signes qui représentent une certaine signification ou un certain événement) et en en déduisant des règles. En réunissant des systèmes de règles complexes, on pourrait déduire logiquement le résultat que l'on veut connaître. Les symbolistes ont donc conçu leurs algorithmes de telle sorte qu'ils produisent des règles à partir des données.

Avec le raisonnement symbolique, la déduction agrandit le champ de la connaissance humaine, tandis que l'induction en élève le niveau. L'induction ouvre de nouveaux champs d'exploration, tandis que la déduction les explore.

Des connexions sur le modèle des neurones du cerveau

Les connexionnistes sont peut-être la plus fameuse des cinq tribus. Ils s'efforcent de reproduire les fonctions cérébrales en utilisant de la silicone à la place des neurones. Fondamentalement, chaque neurone (créé en tant qu'algorithme sur le modèle du neurone réel) résout une petite partie du problème, et l'utilisation d'un grand nombre de neurones en parallèle résout le problème de façon globale.

Le recours à la rétropropagation, ou propagation inverse des erreurs, vise à déterminer les conditions dans lesquelles les erreurs sont corrigées dans les réseaux conçus pour ressembler aux réseaux de neurones chez l'être humain en modifiant les poids (le poids d'une entrée étant son degré d'importance dans le résultat) et les biais (propriétés sélectionnées) dans le réseau. L'objectif est de continuer à changer les poids et les biais jusqu'au moment où le résultat correspond à la cible. À ce stade, le neurone artificiel intervient et transmet sa solution au neurone voisin. La solution créée par un seul neurone n'est qu'une partie de la solution globale. Chaque neurone transmet l'information au neurone suivant jusqu'à ce que le groupe de neurones produise un résultat final. Une telle méthode s'est révélée la plus efficace dans des tâches comme la reconnaissance d'objets, la compréhension du langage écrit ou parlé, et le dialogue avec les utilisateurs.

Les algorithmes évolutionnaires qui testent les variations

Les évolutionnaires se réfèrent aux principes de l'évolution pour résoudre les problèmes. En d'autres termes, cette stratégie se fonde sur la survie des éléments les mieux adaptés (l'élimination des solutions qui ne correspondent pas au résultat voulu). Une fonction de pertinence détermine la viabilité de chaque fonction pour la résolution du problème. La méthode de solution utilise une structure arborescente pour chercher la meilleure solution en fonction de la sortie de fonction. À chaque stade de l'évolution, le gagnant doit élaborer les fonctions du stade suivant. L'idée est que le stade suivant sera plus proche de la résolution du problème mais ne suffira peut-être pas encore pour qu'il soit entièrement résolu, ce qui revient à dire qu'un stade supplémentaire sera nécessaire. Cette tribu, en particulier, est fortement tributaire de la récursivité et de langages qui supportent bien la récursivité pour résoudre les problèmes. Cette stratégie a produit un résultat intéressant, les algorithmes qui évoluent : c'est une génération d'algorithmes qui élabore ceux de la génération suivante.

L'inférence bayésienne

Les scientifiques du groupe des bayésiens ont considéré que l'incertitude était l'aspect essentiel sur lequel il fallait se concentrer, et que l'apprentissage n'était pas assuré mais se réalisait sous forme d'actualisation continue des convictions antérieures, qui devenaient ainsi de plus en plus pertinentes. Cette façon de voir les choses a conduit les bayésiens à adopter des méthodes statistiques, en particulier des méthodes dérivées du théorème de Bayes, permettant de calculer les probabilités dans des conditions spécifiques (par exemple, la probabilité de voir apparaître

une carte d'une certaine couleur, la valeur de départ pour une séquence pseudo-aléatoire, tirée d'un jeu de cartes après trois tirages consécutifs ayant chacun donné cette même couleur).

Les systèmes qui apprennent par analogie

Les analogistes utilisent des fonctions noyaux pour reconnaître des tendances dans les données. En reconnaissant la structure d'une série de données entrantes et en la comparant à la structure d'un résultat connu, on peut élaborer la solution à un problème. L'objectif est d'utiliser la similarité pour déterminer la meilleure solution d'un problème. C'est le type de raisonnement qui détermine que l'utilisation d'une solution particulière a été probante dans une situation donnée à un moment donné et que, par conséquent, on doit aussi pouvoir utiliser cette solution pour une série de situations similaires. Les systèmes de recommandation sont un des résultats les plus caractéristiques de cette tribu. Quand vous achetez un produit sur Amazon, par exemple, le système de recommandation fait apparaître d'autres produits en lien avec ce produit, que vous pourriez avoir également envie d'acheter.

Le but ultime de l'apprentissage machine est d'associer les technologies et les stratégies que recouvrent les cinq tribus de manière à créer un unique algorithme (l'algorithme maître) qui soit capable d'apprendre quelque chose. Naturellement, nous sommes loin d'avoir atteint ce but. Toutefois, des scientifiques comme Pedro Domingos (`http://homes.cs.washington.edu/~pedrod/`) y travaillent en ce moment.

Étudier les trois approches les plus prometteuses en matière d'apprentissage machine

Les dernières sections de ce chapitre étudient les principaux algorithmes choisis par les bayésiens, les symbolistes et les connexionnistes. Ces tribus représentent la frontière présente et future de l'apprentissage à partir des données, car tout progrès vers une IA plus proche de l'être humain dérive de leurs travaux, du moins jusqu'à ce que se produise une nouvelle percée grâce à de nouveaux algorithmes d'apprentissage plus puissants et plus étonnants. Le paysage de l'apprentissage machine est certainement loin de se limiter à ces trois algorithmes, mais si ce chapitre porte sur ces trois tribus, c'est en raison de leur rôle actuel dans l'IA. Voici un résumé des approches abordées dans ce chapitre :

- **Naive Bayes** : cet algorithme peut diagnostiquer certaines maladies avec plus de pertinence et de précision qu'un médecin. En outre, il peut détecter les courriers indésirables et prédire un sentiment à partir d'un texte. Il est aussi largement utilisé dans l'industrie de l'Internet pour traiter plus facilement de gros volumes de données.
- **Les réseaux bayésiens (sous forme graphique)** : ce type de graphe est une représentation de la complexité du monde en termes de probabilités.
- **Les arbres de décision :** l'arbre de décision est ce qui représente le mieux les symbolistes. Il a une longue histoire, et il indique la manière dont une IA peut prendre des décisions car il ressemble à une série de décisions imbriquées.

Le Chapitre 11, « Améliorer l'IA grâce à l'apprentissage profond », est une introduction aux réseaux de neurones, un type d'algorithme proposé par les connexionnistes et qui est le véritable moteur de la renaissance de l'IA. Il traite d'abord du fonctionnement d'un réseau neuronal, et explique ensuite l'apprentissage profond et la raison pour laquelle il est si efficace.



Toutes ces sections sont consacrées à des types d'algorithmes, lesquels sont divisés en sous-catégories. Les arbres de décision, par exemple, se subdivisent en arbres de régression, arbres de classification, le boosting d'arbres, l'ensachage, et la forêt de rotation. On peut même distinguer des sous-types des sous-catégories. Un classifieur de forêts aléatoires est un type d'ensachage, et l'on peut aller plus loin encore dans la subdivision. Au-delà, on commence à voir les vrais algorithmes, qui se comptent par milliers. En résumé, ce livre présente un aperçu d'un sujet infiniment plus complexe, dont la couverture en détail demanderait plusieurs volumes.

Attendre la prochaine percée

Dans les années 1980, quand les systèmes experts dominaient la scène, la plupart des scientifiques et autres professionnels considéraient l'apprentissage machine comme une branche mineure de l'IA servant à apprendre comment répondre au mieux à des prédictions simples provenant de l'environnement (représenté par des données) en utilisant l'optimisation. Aujourd'hui, c'est l'apprentissage machine qui domine et qui l'emporte sur les systèmes experts dans de nombreuses applications et recherches, et c'est sur l'apprentissage machine que reposent des applications que les scientifiques considéraient auparavant comme impossibles à ce niveau de précision et de performance. Ce sont les réseaux de neurones, la solution proposée par les connexionnistes,

qui ont rendu possible cette percée depuis quelques années grâce à des capacités informatiques accrues et à des données plus pertinentes, et grâce aux efforts de chercheurs comme Geoffrey Hinton, Yann LeCun, Yoshua Bengio et bien d'autres.

Les capacités permises par les algorithmes de réseaux de neurones (ce qu'on appelle depuis peu l'apprentissage profond, en raison d'une complexité accrue) augmentent jour après jour. La presse rapporte fréquemment de nouvelles réalisations dans les domaines de la compréhension audio, de la reconnaissance d'image et de vidéo, de la traduction, et même, de la lecture sur les lèvres (si l'apprentissage profond n'est pas encore aussi performant que HAL9000, il n'est cependant pas loin de faire aussi bien que l'être humain : lire l'article à l'adresse https://www.theverge.com/2016/11/7/13551210/ai-deep-learning-lipreading-accuracy-oxford). Ces progrès sont le résultat d'un financement soutenu de la recherche de la part d'entreprises de tailles diverses et de la disponibilité de logiciels performants comme TensorFlow, de Google (https://www.tensorflow.org/) et Computational Network Toolkit (CNTK), de Microsoft qui donnent aux scientifiques et autres professionnels l'accès à la technologie.

On peut s'attendre à des innovations plus sensationnelles encore dans un proche avenir. Bien sûr, il est toujours possible que les chercheurs se heurtent à un mur, comme cela s'est déjà produit lors des précédents hivers de l'IA. Personne ne peut savoir si l'IA atteindra un niveau de capacité comparable à celui de l'être humain en utilisant la technologie actuelle ou si quelqu'un découvrira un algorithme maître, comme le prédit Pedro Domingos (voir https://www.youtube.com/watch?v=qIZ5PXLVZfo), qui résoudrait tous les problèmes d'IA (dont certains qu'il nous reste encore à imaginer). Toujours

est-il que l'apprentissage machine n'est certainement pas une mode encouragée par les médias : il est là pour perdurer, soit dans sa forme actuelle, améliorée, soit sous forme de nouveaux algorithmes à venir.

Explorer la vérité dans les probabilités

Certains sites Internet pourraient vous faire croire que les statistiques et l'apprentissage machine sont deux technologies complètement différentes. La lecture du blog *Statistics vs. Machine Learning, fight !* (http://brenocon.com/blog/2008/12/statistics-vs-machine-learning-fight/), par exemple, laisse penser que non seulement ces deux technologies sont différentes, mais qu'elles sont carrément antagonistes. Si les statistiques constituent une approche plus théorique des problèmes, tandis que l'apprentissage machine est purement fondé sur les données, les statistiques et l'apprentissage machine ont beaucoup de choses en commun. Par ailleurs, les statistiques représentent une des cinq tribus (ou écoles de pensée) grâce auxquelles l'apprentissage machine est devenu viable.

Les statistiques utilisent souvent les probabilités, qui sont un moyen d'exprimer l'incertitude vis-à-vis des événements, et c'est également le cas de l'apprentissage machine et de l'IA (dans une plus large mesure que les statistiques pures). Tous les problèmes ne sont pas similaires à une partie d'échecs ou de Go, dans laquelle vous avez le choix entre un nombre d'actions à la fois vaste et limité. Si vous voulez apprendre à faire évoluer un robot dans un corridor bondé de monde ou à faire en sorte qu'une voiture sans conducteur traverse sans encombre un carrefour, vous devez prendre en compte le fait qu'un

plan (comme par exemple le déplacement d'un point A à un point B) ne comporte pas toujours un résultat unique, mais que plusieurs résultats sont possibles, avec des probabilités différentes. D'une certaine manière, les systèmes d'IA reposent sur un raisonnement fondé sur les probabilités, qui constitue une aide à la décision et permet les choix qui semblent être les meilleurs et les plus rationnels malgré l'incertitude. L'incertitude peut exister pour diverses raisons, et l'IA doit tenir compte du niveau d'incertitude en recourant de façon effective aux probabilités :

1. **Certaines situations n'offrent pas de certitude car elles sont aléatoires par nature. Ce sont des situations intrinsèquement stochastiques. Dans les jeux de cartes, par exemple, on ne peut pas être sûr de la main qu'on obtiendra quand le croupier aura mélangé et distribué les cartes.**

2. **Même lorsqu'une situation n'est pas aléatoire, si l'on n'est pas attentif à tous ses aspects (si l'observation est incomplète), la façon dont les choses évolueront sera incertaine. Ainsi, un robot qui franchit un corridor rempli de monde ne peut pas savoir quelle est la direction que chaque personne veut prendre (ne pouvant pas lire dans leurs pensées), mais il peut formuler une supposition à partir d'une observation partielle du comportement de cette personne. Comme toujours lorsqu'il s'agit d'une supposition, le robot a une certaine chance d'avoir raison et une certaine chance de se tromper.**

3. **Les limites relatives au matériel qui enregistre les données de l'environnement (les capteurs) et les approximations dans le traitement des données peuvent rendre incertains les résultats produits à partir de ces données. Les mesures comportent souvent des erreurs, en raison des**

outils utilisés et de la manière dont elles sont faites. En outre, les êtres humains sont souvent sujets à des biais cognitifs, ils se font facilement des illusions et ils sont parfois aveuglés. De façon similaire, l'IA est limitée par la qualité des données qu'elle reçoit. Les approximations et les erreurs introduisent une incertitude dans chaque algorithme.

Déterminer ce que les probabilités peuvent faire

Une probabilité indique dans quelle mesure un événement a des chances de se produire, et elle s'exprime sous la forme d'un nombre. Si vous lancez une pièce de monnaie en l'air, par exemple, vous ne savez pas si ce sera pile ou face, mais vous pouvez connaître la probabilité des deux résultats. La probabilité d'un événement est comprise entre 0 (aucune possibilité que l'événement se produise) et 1 (certitude que l'événement se produira). Des valeurs intermédiaires comme ¼, ½ et ¾ indiquent que l'événement se produira à une certaine fréquence si l'épreuve susceptible de le déclencher est répétée un nombre de fois suffisamment élevé. En multipliant la probabilité par le nombre d'essais que l'on compte effectuer (un nombre entier), on obtient une estimation du nombre moyen d'occurrences de l'événement si tous ces essais sont effectués. Ainsi, si un événement se produit avec une probabilité p = 0,25 et si l'on effectue 100 essais, l'événement se produira théoriquement 0,25 * 100 = 25 fois.

Il se trouve que p = 0,25 est la probabilité d'obtenir une certaine couleur quand on choisit une carte au hasard dans un jeu de cartes. On explique souvent les probabilités en utilisant l'exemple d'un jeu de 52 cartes.

Les cartes se répartissent à parts égales entre quatre couleurs : les trèfles et les piques, qui sont noirs, et les carreaux et les cœurs, qui sont rouges. Pour déterminer la probabilité de tirer un as, il faut tenir compte du fait qu'il y a quatre as, chacun d'une couleur différente. En termes de probabilité, la réponse est p = 4 / 52 = 0,077.

Une probabilité est nécessairement comprise entre 0 et 1. Les probabilités se définissent empiriquement à partir d'observations. Il suffit de compter le nombre d'occurrences de l'événement en question par rapport à tous les événements qui vous intéressent. Supposons que vous vouliez calculer la probabilité d'une malversation dans les transactions bancaires, ou la fréquence à laquelle les gens attrapent une certaine maladie dans un pays donné. En étant témoin de l'événement, vous pouvez estimer la probabilité qui lui est associée en comptant le nombre de fois que l'événement se produit et en divisant ce nombre par le nombre total d'épreuves.

Vous pouvez compter le nombre d'occurrences de malversations ou d'une maladie en utilisant des données archivées (dans des bases de données) et en divisant ce nombre par le nombre total d'événements génériques ou d'observations. Vous divisez donc le nombre de malversations par le nombre de transactions en une année, ou le nombre de personnes atteintes de la maladie au cours de l'année par la population de la région concernée. Le résultat est un nombre compris entre 0 et 1, qui sera votre probabilité de référence pour un certain événement dans des circonstances données.

Sachant qu'il n'est pas toujours possible de compter les occurrences d'un événement, il faut savoir échantillonner. L'échantillonnage, qui est fondé sur certaines attentes en matière de probabilités, permet, en étudiant seulement une petite partie d'un ensemble d'événements ou d'objets, d'en déduire correctement leurs probabilités ainsi que des mesures quantitatives exactes ou des classes qualitatives

liées à une série d'objets. Par exemple, si vous voulez étudier les ventes de voitures au niveau national au cours du dernier mois, vous n'avez pas besoin de suivre chaque vente. En utilisant un échantillon constitué des ventes d'un petit nombre de concessionnaires répartis dans tout le pays, vous pouvez déterminer des mesures quantitatives comme le prix moyen d'une voiture vendue, ou des mesures qualitatives, comme le modèle de voiture le plus vendu.

Prendre en compte les connaissances préalables

Les probabilités s'entendent en termes de temps et d'espace, mais d'autres conditions influencent aussi la probabilité que l'on mesure. Le contexte est important. Quand on estime la probabilité d'un événement, on peut (parfois à tort) avoir tendance à penser qu'on peut appliquer la probabilité calculée à chaque situation possible. C'est ce que l'on appelle la *probabilité a priori*.

Ainsi, si vous lancez une pièce de monnaie, en supposant qu'elle est bien équilibrée, la probabilité *a priori* d'obtenir face est d'environ 50 % (en supposant aussi l'existence d'une minuscule probabilité que la pièce tombe sur la tranche). Quel que soit le nombre de lancers réalisés, la probabilité d'obtenir face au prochain lancer sera toujours d'environ 50 %. Cependant, dans d'autres situations, si vous changez le contexte, la probabilité *a priori* n'est plus la même. Dans ce cas, on peut parler de *probabilité a posteriori*. Il s'agit de la probabilité *a priori* une fois qu'un événement a modifié la donne.

Par exemple, la probabilité *a priori* qu'une personne soit de sexe féminin est à peu près égale à 50 %, mais cette probabilité peut être nettement différente si l'on ne prend

en compte que les dernières tranches d'âge, sachant que les femmes vivent généralement plus longtemps. Ces tranches d'âge comporteront donc plus de femmes que d'hommes. Un autre exemple lié au sexe est celui des études supérieures, sachant que les femmes sont plus nombreuses que les hommes dans les plus grandes universités.

En termes de répartition des sexes, la nature et la culture peuvent l'une et l'autre engendrer une probabilité *a posteriori* différente de la probabilité *a priori*. Les sections qui suivent vous permettent de comprendre plus en détail l'utilité des probabilités.

Probabilités conditionnelles et Naive Bayes

Dans des situations comme celles mentionnées dans la section précédente, relatives à la répartition des sexes, on parle de probabilité conditionnelle, notée p (y|x), soit « p de y sachant x », ce qui signifie : probabilité que se produise l'événement y sachant que l'événement x s'est produit. Les probabilités conditionnelles sont un outil très puissant pour l'apprentissage machine et l'IA. En fait, si la probabilité *a priori* peut changer autant en raison de certaines circonstances, connaître les circonstances possibles augmente les chances de prédire correctement un événement en observant des exemples : c'est précisément à cela que sert l'apprentissage machine. Ainsi, comme on l'a vu précédemment, la probabilité qu'une personne prise au hasard soit de sexe masculin, ou féminin, est généralement de 50 %, mais que se passe-t-il si l'on sait que cette personne a les cheveux longs, ou les cheveux courts ? On peut estimer la probabilité d'avoir les cheveux longs à 35 % de la population ; cependant, si l'on ne s'intéresse qu'à la population féminine, cette probabilité est de 60 %. Si ce

pourcentage est si élevé dans la population féminine, contrairement à la probabilité *a priori*, l'algorithme d'apprentissage machine appelé Naive Bayes a besoin d'une donnée qui lui indique si les cheveux de la personne sont longs ou courts.

En fait, l'algorithme Naive Bayes augmente les chances de prédiction correcte en connaissant les circonstances dans lesquelles se fait la prédiction. Tout commence avec le révérend Thomas Bayes et son théorème révolutionnaire sur les probabilités. Comme on l'a noté ailleurs dans ce livre, une des tribus de l'apprentissage machine porte son nom (les bayésiens). Les bayésiens utilisent diverses méthodes statistiques pour résoudre les problèmes, et toutes ces méthodes consistent à observer les probabilités du résultat désiré dans le contexte approprié, avant et après avoir observé le résultat lui-même. À partir de ces observations, ils résolvent le problème du lever du soleil (consistant à estimer la probabilité que le soleil se lève demain) en enchaînant des observations répétées et en actualisant continuellement leur estimation de la probabilité que le soleil se lève de nouveau par une proportion du nombre de fois où ils ont déjà vu le soleil se lever. À propos de l'application du raisonnement de Bayes à un bébé qui observe le soleil, lisez cet article du magazine *The Economist* à l'adresse http://www.economist.com/node/382968.

Les spécialistes de la science des données attendent beaucoup du développement des algorithmes avancés basés sur les probabilités de Bayes. Le magazine du MIT *Technology Review* mentionne l'apprentissage machine bayésien comme étant une technologie émergente qui changera notre monde (http://www2.technologyreview.com/news/401775/10-emerging-technologiesthat-will-change-the/). Néanmoins, les fondements du théorème de Bayes ne sont pas si compliqués (ils peuvent

simplement être quelque peu contrintuitifs si l'on ne tient compte que des probabilités *a priori*, comme la plupart des gens, et non des probabilités *a posteriori*).

Étudier le théorème de Bayes

Pasteur presbytérien, le révérend Thomas Bayes était aussi statisticien et philosophe. Il a formulé son théorème au cours de la première moitié du XVIIIe siècle. Ce théorème n'a jamais été publié de son vivant. Sa publication a révolutionné la théorie des probabilités en introduisant la notion de probabilité conditionnelle (mentionnée dans la section précédente). Grâce au théorème de Bayes, prédire la probabilité qu'une personne soit un homme ou une femme devient plus facile si l'on sait que cette personne a des cheveux longs. La formule utilisée par Thomas Bayes est la suivante :

P(B|E) = P(E|B) * P(B) / P(E)



Le révérend Bayes n'a pas conçu l'algorithme du classifieur bayésien naïf (Naive Bayes) : il a seulement formulé le théorème. En fait, on ne sait pas vraiment à qui attribuer cet algorithme. Il est apparu pour la première fois dans un manuel en 1973 sans aucune référence à son créateur et il est passé inaperçu pendant plus d'une décennie, jusqu'à ce qu'en 1990, des chercheurs remarquent qu'il produisait des prédictions remarquablement précises lorsqu'il était alimenté avec suffisamment de données exactes. Lire la formule en utilisant comme entrée l'exemple précédent permet de mieux la comprendre, sans quoi elle est plutôt contrintuitive :

- **P(B|E) :** probabilité d'un événement (B) compte tenu d'un fait connu (E) (probabilité postérieure). Il convient de voir en B un autre moyen d'exprimer une hypothèse. En l'occurrence, l'hypothèse est qu'une

personne est de sexe féminin, et le fait connu est qu'elle a des cheveux longs. Connaître la probabilité associée à cette hypothèse, compte tenu de ce qui est connu, permet de prédire le sexe de la personne avec un certain niveau de confiance.

» **P(E|B) :** probabilité qu'une personne ait des cheveux longs sachant qu'elle est de sexe féminin. Ce terme fait référence à la probabilité du fait connu dans le sous-groupe, laquelle est une probabilité conditionnelle. En l'occurrence, cette probabilité est de 60 %, soit la valeur 0,6 dans la formule (probabilité antérieure).

» **P(B) :** probabilité générale que la personne soit de sexe féminin, soit la probabilité *a priori* de l'événement hypothétique. En l'occurrence, cette probabilité est de 50 %, soit la valeur 0,5 dans la formule.

» **P(E) :** probabilité générale que la personne ait des cheveux longs. C'est là encore une probabilité *a priori*, liée cette fois à l'événement observé. En l'occurrence, cette probabilité est de 35 %, soit la valeur 0,35.

Quand on résout le problème qui précède à l'aide de la formule de Bayes et des valeurs de probabilité que l'on a déterminées, le résultat est : 0,6 * 0,5 / 0,35 = 0,857. C'est une probabilité élevée, ce qui conduit à affirmer que compte tenu de ce que l'on sait, la personne en question est probablement de sexe féminin.

Un autre exemple courant, qui peut surprendre et que l'on rencontre souvent dans les manuels et les magazines scientifiques, est celui du test médical positif, qui est intéressant pour mieux comprendre comment les probabilités antérieures et postérieures peuvent changer, notablement quand les circonstances sont différentes.

Supposons que vous soyez préoccupé par l'idée d'être potentiellement atteint d'une maladie rare qui touche 1 % de la population. Vous avez subi un test dont le résultat est positif. Les tests médicaux ne sont jamais d'une fiabilité parfaite, et le laboratoire vous dit que si vous êtes effectivement atteint, le test sera positif dans 99 % des cas, tandis que si vous êtes en bonne santé, le test sera négatif dans 99 % des cas. Compte tenu de ces chiffres, vous pensez tout de suite que vous êtes atteint par la maladie. Or, la réalité est très différente. En l'occurrence, le théorème de Bayes s'applique aux valeurs suivantes :

- 0,99 pour P(E|B)
- 0,01 pour P(B)
- 0,01 * 0,99 + 0,99 * 0,01 = 0,0198 pour P(E)

On calcule alors 0,01 * 0,99 / 0,0198 = 0,5, soit une probabilité de 50 % que vous soyez malade. Au final, vos chances de ne pas l'être sont plus grandes que ce que vous pensiez. Vous vous demandez peut-être comment c'est possible. En termes de réponse positive au test, les proportions sont les suivantes :

- **Les personnes qui sont malades et pour qui le test donne la réponse correcte :** ce groupe est celui des vrais positifs, et il représente 99 % du 1 % de la population ayant la maladie.
- **Les personnes qui ne sont pas malades et pour qui le test ne donne pas la réponse correcte :** ce groupe est constitué du 1 % obtenant une réponse positive au test parmi les 99 % de la population qui ne sont pas malades. Ici encore, on multiplie 99 % par 1 %. C'est le groupe des faux positifs.

Quand on étudie le problème sous cet angle, les choses deviennent évidentes. Si l'on se limite aux personnes qui obtiennent une réponse positive au test, la probabilité

qu'une personne se trouve dans le groupe des vrais positifs est égale à la probabilité qu'elle se trouve dans le groupe des faux positifs.

Se représenter le monde sous la forme d'un graphe

Le théorème de Bayes vous permet de déduire la probabilité qu'un événement se produise dans un certain contexte, à partir des probabilités générales de l'événement lui-même et du fait connu, combinées à la probabilité du fait connu sachant que l'événement se produit. Il est rare qu'un seul fait connu mette fin aux doutes en introduisant dans une prédiction suffisamment de certitude pour que l'on sache que l'événement va se produire. Comme un vrai détective, pour aboutir à la certitude, vous devez recueillir davantage de preuves et assembler les pièces du puzzle. Remarquer qu'une personne a des cheveux longs ne suffit pas pour déterminer si c'est un homme ou une femme. Des données supplémentaires sur sa taille et son poids permettraient de le prédire avec un degré de confiance nettement plus élevé.

L'algorithme Naive Bayes vous permet de réunir toutes les données disponibles et d'obtenir une prédiction plus solide avec une probabilité plus élevée qu'elle soit correcte. Considérées isolément, les données connues n'éliminent pas le risque d'une prédiction incorrecte, mais prises ensemble, elles peuvent permettre d'aboutir à une solution plus définitive. L'exemple suivant est l'illustration d'une classification par l'algorithme Naive Bayes. Il s'agit d'un problème ancien et bien connu, mais qui illustre bien les possibilités que l'on peut attendre d'une IA. Les données sont tirées de l'article « Induction of Decision Trees » de John Ross Quinlan

(http://dl.acm.org/citation.cfm?id=637969). Quinlan est un chercheur en informatique qui a contribué au développement des algorithmes d'apprentissage machine sur les arbres de décision, mais son exemple fonctionne avec n'importe quel algorithme d'apprentissage. L'IA doit prédire les meilleures conditions pour jouer au tennis compte tenu de la situation météorologique. L'inventaire des propriétés retenu par Quinlan est le suivant :

» Temps : ensoleillé, couvert ou pluvieux

» Température : froid, doux ou chaud

» Humidité : élevée ou normale

» Venté : vrai ou faux

Le tableau suivant contient les entrées de la base de données utilisées dans cet exemple :

Temps Ensoleillé	Température Chaud	Humidité Élevée	Venté Faux	JouerAuTennis Non
Ensoleillé	Chaud	Élevée	Vrai	Non
Couvert	Chaud	Élevée	Faux	Oui
Pluvieux	Doux	Élevée	Faux	Oui
Pluvieux	Froid	Normale	Faux	Oui
Pluvieux	Doux	Élevée	Vrai	Non
Pluvieux	Froid	Normale	Vrai	Non
Couvert	Froid	Normale	Vrai	Oui
Ensoleillé	Doux	Élevée	Faux	Non
Ensoleillé	Froid	Normale	Faux	Oui
Pluvieux	Doux	Normale	Faux	Oui
Ensoleillé	Doux	Normale	Vrai	Oui

Couvert	Doux	Élevée	Vrai	Oui
Couvert	Chaud	Normale	Faux	Oui

Le choix de jouer au tennis dépend des quatre arguments de la Figure 10.1.

Le résultat de cet exemple d'apprentissage de l'IA est la décision de jouer ou de ne pas jouer au tennis, compte tenu des conditions météorologiques (données connues). Le seul critère du temps (ensoleillé, couvert ou pluvieux) ne suffirait pas, car la température et l'humidité pourraient être trop élevées, ou bien le vent pourrait être trop fort. Ces arguments représentent des conditions réelles dont les causes sont multiples ou interdépendantes. L'algorithme Naive Bayes sait prédire de façon performante quand les causes sont multiples.

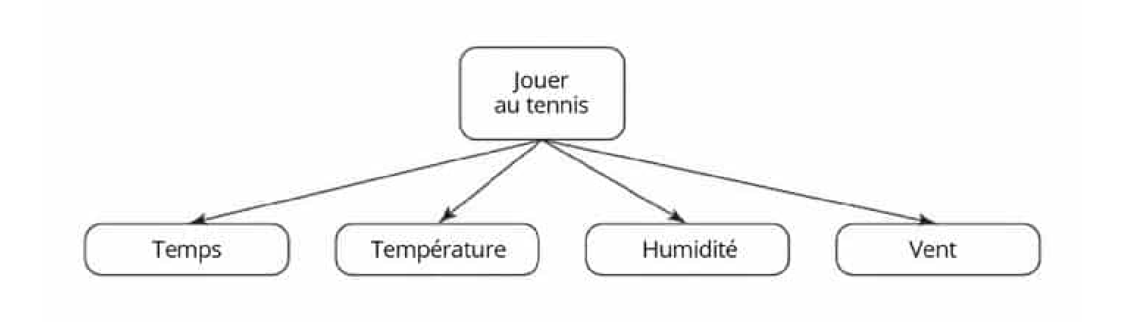


FIGURE 10.1 : Un modèle bayésien naïf peut reconstituer les faits qui déterminent le résultat correct.

L'algorithme calcule un score en fonction de la probabilité de prendre une décision particulière, et le multiplie par les probabilités des données connues liées à cette décision. Ainsi, pour déterminer si l'on peut jouer au tennis quand le temps est ensoleillé et le vent est fort, l'algorithme calcule le score pour une réponse positive en multipliant la probabilité générale de jouer (9 cas sur 14 occurrences) par la probabilité que la journée soit ensoleillée (2 cas

sur 9 situations dans lesquelles on joue au tennis) et qu'il vente alors qu'on joue au tennis (3 cas sur 9 situations dans lesquelles on joue au tennis).

Les mêmes règles s'appliquent aux situations négatives (avec des probabilités différentes de ne pas jouer au tennis compte tenu de certaines conditions climatiques) :

```
Probabilité de jouer : 9 / 14 * 2 / 9 * 3 / 9 = 0,05
Probabilité de ne pas jouer : 5 / 14 * 3 / 5 * 3 / 5 = 0,13
```

Le score étant élevé, l'algorithme décide qu'il vaut mieux ne pas jouer dans de telles conditions. Il calcule les probabilités en additionnant les deux scores et en les divisant par leur somme :

```
Probabilité de jouer : 0,05 / (0,05 + 0,13) = 0,278
Probabilité de ne pas jouer : 0,13 / (0,05 + 0,13) = 0,722
```

Avec l'algorithme Naive Bayes, en utilisant un *réseau bayésien* constitué de graphes qui montrent l'interdépendance entre les événements, on peut aussi représenter des relations plus complexes qu'une série de facteurs permettant de calculer la probabilité d'un résultat. Les graphes bayésiens comportent des nœuds qui représentent les événements, et des arcs qui montrent quelles sont les influences des différents événements sur les autres, et s'accompagnent d'un tableau de probabilités conditionnelles montrant comment fonctionnent les relations en termes de probabilités. La Figure 10.2 montre un exemple fameux de réseau bayésien tiré d'un article académique de 1988, « Local computations with probabilities on graphical structures and their application to expert systems », de Steffen L. Lauritzen et David J. Spiegelhalter, publié dans le *Journal of the Royal Statistical Society* (voir https://www.jstor.org/stable/2345762).

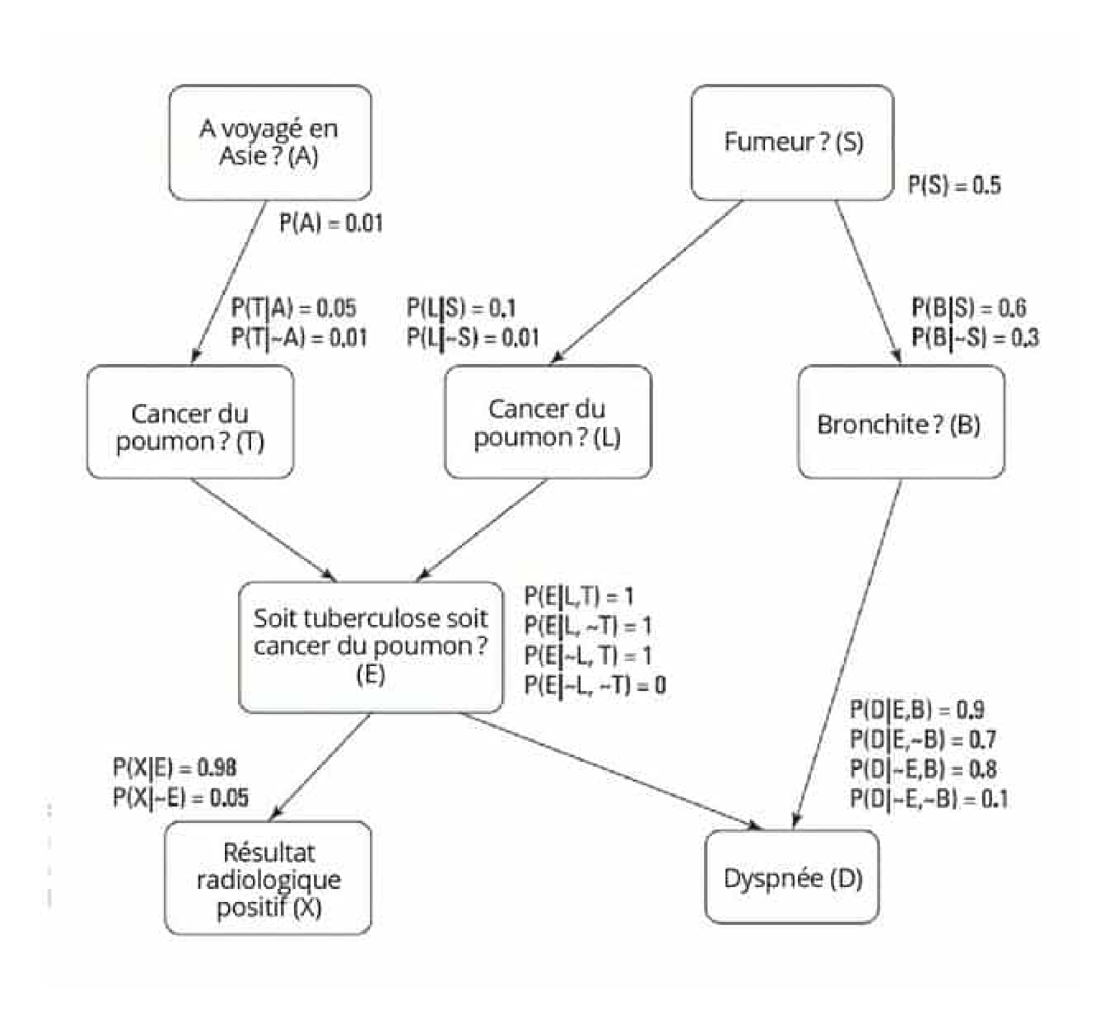


FIGURE 10.2 : Un réseau bayésien peut servir d'aide à la décision dans le domaine médical.

Le réseau présenté ci-dessus s'appelle *Asia*. Il représente plusieurs états pathologiques possibles et leurs causes. Si un patient est atteint de dyspnée, par exemple, ce peut être l'effet d'une tuberculose, d'un cancer du poumon ou d'une bronchite. Le fait de savoir si le patient fume, s'il est allé en Asie et si ses examens radiologiques présentent des anomalies (autant d'éléments de certitude, ce qu'on appelle l'*a priori* dans le langage bayésien) permet de déduire les probabilités réelles (postérieures) qu'il souffre de telle ou telle pathologie représentée sur le graphe.

Les réseaux bayésiens, même s'ils ont un aspect intuitif, reposent sur des mathématiques élaborées et font mieux

qu'un simple modèle bayésien naïf, car ils représentent le monde sous forme d'une séquence de causes et d'effets basée sur les probabilités. Les réseaux bayésiens peuvent être utilisés pour représenter n'importe quelle situation. Ils trouvent des applications variées : diagnostics médicaux, fusion des données relatives à l'incertain en provenance de capteurs multiples, modélisation économique, et gestion de systèmes complexes, notamment dans une automobile. Ainsi, sachant que la circulation sur voie rapide peut mettre en jeu des situations complexes impliquant un certain nombre de véhicules, le consortium Analysis of MassIve Data STreams (AMIDST), en collaboration avec le motoriste Daimler, a conçu un réseau bayésien capable d'identifier les manœuvres effectuées par les autres véhicules et de faire ainsi progresser la sécurité routière.

Faire grandir des arbres pour classer des éléments

L'arbre de décision est un autre type d'algorithme essentiel dans le domaine de l'apprentissage machine et de la mise en œuvre de l'IA. Les algorithmes sous forme d'arbres de décision ne sont pas une nouveauté, ils ont une longue histoire. Le premier algorithme de ce type date des années 1970 (il y a eu par la suite de nombreuses variantes). Si l'on prend en compte les expérimentations et les premières recherches, l'utilisation d'arbres de décision remonte plus loin encore : aussi loin que les perceptions. En tant que principaux algorithmes de type symboliste, les arbres de décision connaissent depuis longtemps un grand succès, de par leur caractère intuitif. Il est facile de traduire leur résultat sous forme de règles, et de le rendre ainsi très compréhensible pour les humains. Les arbres de décision sont aussi extrêmement faciles à utiliser. Toutes

ces propriétés en font un outil simple, attrayant et efficace pour les modèles qui requièrent des transformations complexes de matrices de données d'entrée ou un réglage extrêmement précis des hyperparamètres.



Le symbolisme est l'approche de l'IA fondée sur les énoncés logiques et le recours généralisé à la déduction. La *déduction* enrichit la connaissance à partir de ce qui est connu, et l'*induction* formule des règles générales sur la base des données et faits connus.

Prédire les résultats en fractionnant les données

Si vous disposez d'une série de mesures et si vous voulez les décrire en utilisant une valeur unique, ce sera la *moyenne* arithmétique (calculée en additionnant toutes les mesures et en divisant le nombre obtenu par le nombre de mesures). De façon similaire, si vous disposez d'une série de classes ou de qualités (s'il s'agit par exemple d'une série de données relatives à différentes races de chiens ou à différents types de produits), vous pouvez vous servir de la classe la plus importante ou du groupe le plus important, c'est-à-dire du *mode*, pour représenter l'ensemble de la série. Le mode, comme la moyenne, est une mesure statistique. Il est constitué de la valeur (une mesure ou une classe) qui apparaît le plus souvent. La moyenne et le mode consistent à présenter une valeur ou une classe qui prédit le prochain élément en minimisant l'erreur. D'une certaine manière, ce sont des indicateurs qui apprennent la réponse à partir des données existantes. Avec les arbres de décision, les moyennes et les modes sont des indicateurs, sachant que le jeu de données est divisé en séries de données de moindre dimension dont les moyennes ou les modes sont les meilleurs indicateurs possibles pour le problème étudié.



Diviser un problème afin de parvenir plus facilement à une solution est aussi une stratégie courante utilisée par un certain nombre d’algorithmes de type *diviser pour régner*. Au cours d’un combat, on peut plus facilement être victorieux en divisant l’armée ennemie de manière à mener des combats singuliers. À partir d’un échantillon d’observations, l’algorithme reconstitue les règles qui ont généré les classes obtenues en sortie (ou les valeurs numériques, s’il s’agit d’un problème de régression) en divisant la matrice d’entrée en partitions de plus en plus petites jusqu’à ce que le processus déclenche une règle d’arrêt. Cette manière de progresser du particulier vers le général est typique de la déduction inverse, telle qu’elle est étudiée dans le domaine de la logique et de la philosophie.



Dans le contexte de l’apprentissage machine, ce type de raisonnement inverse consiste à effectuer une recherche de toutes les manières possibles de diviser l’apprentissage et à décider, selon le principe de l’algorithme glouton, de retenir la forme de division qui maximise les mesures statistiques des partitions résultantes. On parle d’algorithme *glouton* à propos d’un algorithme qui prend toujours les décisions qui maximisent le résultat de l’étape en cours du processus d’optimisation, sans tenir compte des conséquences éventuelles au cours des étapes suivantes. Il s’ensuit qu’un algorithme glouton n’assure pas nécessairement l’optimisation au niveau global.

La division est l’application d’un principe simple : chaque partition des données initiales doit faciliter la prédiction du résultat cible, qui se caractérise par une distribution des classes (ou des valeurs) différente et plus favorable que l’échantillon initial. L’algorithme crée des partitions en divisant les données. Il détermine la répartition en évaluant d’abord les propriétés, puis il mesure les valeurs pouvant maximiser une mesure statistique particulière : la

mesure qui joue le rôle de fonction de coût dans un arbre de décision.

Un certain nombre de mesures statistiques déterminent la manière de pratiquer la répartition des données dans l'arbre de décision. La règle est toujours que la division doit constituer une amélioration par rapport à l'échantillon initial et par rapport à toute autre division possible, lorsque cela rend la prédiction plus sûre. On mesure le plus souvent le coefficient d'impureté de Gini, le gain d'information et la réduction de la variance (pour les problèmes de régression). Sachant que ces mesures fonctionnent de façon similaire, ce chapitre est centré sur le gain d'information, qui est la mesure la plus intuitive, et montre comment un arbre de décision peut détecter une capacité prédictive améliorée (ou une réduction de risque) de la façon la plus facile pour un certain fractionnement. L'algorithme d'arbre de décision en fonction du gain d'information (ID3) que Ross Quinlan avait créé dans les années 1970 est toujours utilisé, grâce à sa récente mise à jour C4.5. Le gain d'information suit la formule de l'entropie informative (élaborée par Claude Shannon, un mathématicien et ingénieur américain considéré comme le père de la théorie de l'information), une formule à caractère général qui exprime la valeur attendue de l'information contenue dans un message :

```
Entropie de Shannon E = -∑(p(i)X~log2(p(i)))
```

Dans cette formule, on passe en revue toutes les classes, une par une, et l'on additionne les résultats de la multiplication. Dans la multiplication effectuée au niveau de chaque classe, p(i) est la probabilité associée (exprimée dans l'intervalle compris entre 0 et 1) et log2 est le logarithme de base 2. En partant d'un échantillon dans lequel on veut classer deux intervalles ou deux classes ayant la même probabilité (distribution 50/50), l'entropie maximale est : Entropie = - 0,5 * log2 (0,5) -0,5 * log2 (0,5)

= 1,0. Cependant, quand l'algorithme associé à l'arbre de décision détecte une propriété à partir de laquelle on peut diviser la série de données en deux partitions avec une distribution 40/60 entre les deux classes, l'entropie informative moyenne diminue :

```
Entropie = -0,4 * log2(0,4) -0,6 * log2(0,6) = 0,97
```

On note la somme entropique pour l'ensemble des classes. Avec une répartition 40/60, la somme est moins élevée que le maximum théorique de 1 (diminution de l'entropie). L'entropie peut être considérée comme une mesure du désordre dans les données : moins il y a de désordre, plus il y a d'ordre et plus il est facile de prédire la bonne classe. Après une première répartition, l'algorithme divise encore les partitions obtenues selon la même logique de réduction de l'entropie. Il divise progressivement les partitions de données jusqu'à ce qu'il n'y ait plus de division possible parce que le sous-échantillon obtenu est un élément unique, ou parce qu'une règle d'arrêt a été atteinte.

Les *règles d'arrêt* sont les limites de l'expansion d'un arbre. Elles suivent la prise en compte de trois aspects de la partition : la taille de la partition initiale, la taille de la partition résultante, et le gain d'information que la division permet d'obtenir. Elles sont importantes parce que les algorithmes d'arbres de décision approchent un grand nombre de fonctions. Toutefois, le bruit et les erreurs de données peuvent facilement influencer cet algorithme. En conséquence, en fonction de l'échantillon, l'instabilité et la variance des estimations résultantes ont une incidence sur les prédictions de l'arbre de décision.

Prendre des décisions en fonction des arbres

Dans cette section, l'exemple utilisé est celui de la série de données de Ross Quinlan dont il a été question dans la section « Se représenter le monde sous la forme d'un graphe », précédemment dans ce chapitre. Cela nous permet de présenter l'algorithme ID3, un arbre de décision particulier que l'on trouve dans l'article « Induction of Decision Trees » mentionné précédemment. La série de données est très simple, elle n'est constituée que de 14 observations relatives à la météorologie, avec des résultats qui indiquent s'il est approprié ou non de jouer au tennis.

Cet exemple comprend quatre propriétés : le temps, la température, l'humidité et le vent. Elles sont exprimées sous forme de classes qualitatives plutôt que de mesures quantitatives (la température, l'humidité et la force du vent peuvent aussi être exprimées sous forme numérique) pour permettre de comprendre de façon plus intuitive le lien entre la météo et le résultat. Quand ces propriétés sont traitées par l'algorithme, on peut représenter la série de données sous forme d'un schéma arborescent, comme le montre la Figure 10.3. On peut inspecter et lire une série de règles en fractionnant la série de données de manière à obtenir différentes parties dans lesquelles les prédictions seront plus faciles, sachant qu'il suffira de voir quelle est la classe la plus fréquente (en l'occurrence, le résultat est la décision de jouer ou de ne pas jouer au tennis).

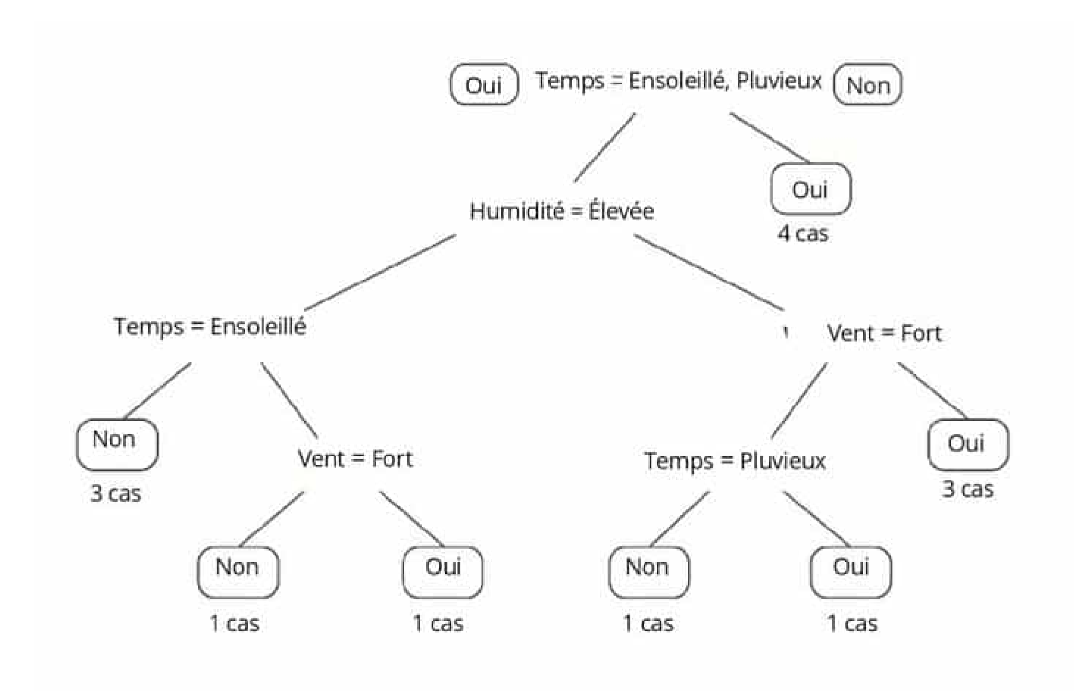


FIGURE 10.3 : Une visualisation de l'arbre de décision construit à partir des données concernant la possibilité de jouer au tennis.

Pour lire les nœuds de l'arbre, partez du nœud supérieur, qui correspond aux données initiales de l'apprentissage, puis lisez les règles. Notez que chaque nœud possède deux dérivations : la branche de gauche signifie que la règle inscrite au-dessus est vraie (le mot « oui » est entouré) et la branche de droite signifie que cette règle est fausse (le mot « non » est entouré).

À droite de la première règle, on voit une importante règle terminale (une feuille terminale), dans un cercle, indiquant un résultat positif, « Oui », qui signifie que Jouer au tennis = Vrai. D'après ce nœud, quand le temps n'est ni ensoleillé (Soleil) ni pluvieux (Pluie), il est possible de jouer (le nombre sous la feuille terminale indique que quatre exemples confirment cette règle et qu'aucun ne l'infirme). On peut noter que cette règle serait plus claire s'il était simplement énoncé qu'il est possible de jouer quand le temps est couvert. Il arrive souvent que les règles des arbres de décision ne soient pas

immédiatement utilisables, et qu'il faille tout d'abord les interpréter. Cependant, elles sont clairement intelligibles (et elles le sont bien davantage qu'un vecteur de valeurs).

Sur la gauche, l'arbre continue avec d'autres règles concernant l'humidité. À gauche toujours, quand l'humidité est élevée et quand le temps est ensoleillé, les feuilles terminales sont négatives, sauf lorsque le vent est fort. Quand on explore les branches sur la droite, l'arbre révèle qu'il n'est pas toujours possible de jouer quand le vent n'est pas fort, ni quand le vent est fort mais qu'il ne pleut pas.

Élaguer les arbres qui ont bien poussé

Si le jeu de données de la section précédente, concernant la possibilité de jouer au tennis, illustre les aspects qui caractérisent un arbre de décision, il est cependant peu intéressant d'un point de vue probabiliste car il propose une série d'actions déterministes (en l'absence d'instructions conflictuelles). Un apprentissage avec des données réelles ne présente généralement pas des règles aussi précises, il laisse place à l'ambiguïté et à la possibilité du résultat espéré.

Les arbres de décision présentent davantage de variance que de biais dans leurs estimations. Pour limiter le surapprentissage, il est spécifié dans l'exemple que le fractionnement minimum doit comporter au moins cinq exemples ; l'arbre est aussi élagué. L'élagage se produit quand l'arbre est bien développé.

À partir des feuilles, l'arbre est élagué de ses branches, mais cela n'améliore que peu la réduction du gain d'information. Quand on laisse d'abord l'arbre se développer, les branches qui offrent peu d'amélioration

sont tolérées parce qu'elles débouchent sur des branches et des feuilles plus intéressantes. Le fait de remonter des branches vers la racine en ne gardant que les branches qui ont une valeur prédictive réduit la variance du modèle, et les règles résultantes deviennent parcimonieuses.



Dans le cas d'un arbre de décision, l'élagage est similaire au brainstorming. Tout d'abord, le code génère toutes les ramifications possibles (comme on exprime toutes les idées possibles durant une séance de brainstorming). Ensuite, quand le brainstorming est terminé, le code ne conserve que ce qui fonctionne vraiment.

Chapitre 11

Améliorer l’IA grâce à l’apprentissage profond

DANS CE CHAPITRE

» Commencer avec le perceptron limité

» Maîtriser les éléments du réseau de neurones et de la rétropropagation

» Percevoir et détecter des objets dans des images à l’aide de convolutions

» Utiliser des séquences et les détecter avec des réseaux de neurones récurrents

» Découvrir l’aspect créatif de l’IA grâce aux réseaux antagonistes génératifs

Les journaux, les magazines, les réseaux sociaux et les sites Internet disent tous la même chose : l’IA est une innovation intéressante qui va révolutionner le monde grâce à l’apprentissage profond. Or, l’IA recouvre bien davantage que l’apprentissage machine, dont l’apprentissage profond ne représente qu’une petite partie.

Il est important de faire la distinction entre ce qui est vanté pour appâter les investisseurs et ce que cette technologie peut réellement faire, et c’est l’objectif global de ce chapitre.

Ce chapitre vous présente l’apprentissage profond d’un point de vue pratique et technique et vous montre ce qu’il peut offrir à court terme, en explorant ses possibilités et

ses limites. Le chapitre commence par un historique des réseaux de neurones, dont il précise les bases. Il présente ensuite les résultats incomparables des réseaux de neurones convolutifs, des réseaux de neurones récurrents (ces deux types de réseaux étant destinés à l'apprentissage machine supervisé) et des réseaux antagonistes génératifs (une sorte d'apprentissage machine non supervisé).

Développer des réseaux de neurones similaires au cerveau humain

Les sections qui suivent présentent une famille d'algorithmes d'apprentissage qui s'inspirent du fonctionnement du cerveau. Ce sont les réseaux de neurones, le principal algorithme de la tribu des connexionnistes qui imite le mieux les neurones du cerveau humain, sur une plus petite échelle.



Le connexionnisme est l'approche de l'apprentissage machine fondée sur les neurosciences et sur l'exemple des réseaux biologiquement interconnectés.

Faire entrer en jeu le neurone

Le cerveau humain est constitué de plusieurs millions de neurones, c'est-à-dire de cellules qui reçoivent, traitent et transmettent des signaux électriques et chimiques. Chaque neurone possède un noyau et des filaments qui servent de transducteurs d'entrée, des dendrites qui reçoivent les signaux provenant des autres neurones, et un unique filament de sortie, l'axone, qui se termine par des synapses assurant la communication vers l'extérieur.

Les neurones sont reliés à d'autres neurones et se transmettent l'information par des substances chimiques, tandis que l'information à l'intérieur même du neurone est traitée de façon électrique.

L'ingénierie inverse du traitement du signal par le cerveau permet aux connexionnistes de définir des réseaux neuronaux basés sur une analogie avec la biologie et leurs composants. Des termes sont empruntés à la biologie du cerveau, comme neurones, activation, et connexions, et désignent des opérations mathématiques. Les réseaux neuronaux ressemblent surtout à des séries de multiplications et d'additions. Ce sont cependant des algorithmes extraordinairement efficaces pour résoudre des problèmes complexes comme la reconnaissance d'images et de sons ou la traduction automatique. Avec un matériel spécialisé, ils peuvent exécuter rapidement des calculs de prédiction.

VOIR L'APPRENTISSAGE PROFOND COMME UNE AUGMENTATION

Le Chapitre 10 traite des réseaux bayésiens et présente un exemple de la manière dont ces réseaux peuvent aider le médecin à formuler un diagnostic. Pour ce faire, le réseau bayésien nécessite une préparation adéquate des données de probabilités. L'apprentissage profond permet de créer un lien entre la capacité des algorithmes de prendre la meilleure décision possible, en utilisant toutes les données requises, et les données qui sont réellement disponibles, et qui ne sont jamais dans le format que les algorithmes d'apprentissage machine comprennent. Qu'il s'agisse de photos, d'images, de sons enregistrés, de données de l'Internet (en particulier celles des réseaux sociaux) ou des dossiers d'une entreprise, une analyse de données est toujours nécessaire pour obtenir des données exploitables.

Dans l'avenir, un algorithme d'apprentissage profond pourrait aider les médecins en mettant leurs connaissances approfondies en médecine en correspondance avec les informations sur les patients (en utilisant toutes les sources disponibles, telles que les livres, les livres blancs et les résultats des recherches les plus récentes). Les informations relatives aux patients pourraient elles-mêmes provenir des diagnostics et des prescriptions qui auront précédé, ou même, des informations disponibles sur les réseaux sociaux (ainsi, par exemple, le médecin n'aurait pas besoin de demander à son patient s'il est allé en Asie puisque l'IA aurait détecté et identifié cette information à partir de ses photos publiées sur Instagram ou sur Facebook). Ce scénario peut ressembler à de la science-fiction, mais la création d'un tel système est presque déjà possible : une IA peut dès à présent détecter une pneumonie sur des radios mieux qu'un radiologue saurait le faire, grâce au Stanford Machine Learning Group

(https://stanfordmlgroup.github.io/projects/chexnet/).

L'apprentissage profond apparaît aussi dans un certain nombre d'applications. Il est présent sur les réseaux sociaux là où les images et autres contenus sont classés automatiquement, dans les moteurs de recherche quand ils répondent aux requêtes, dans les publicités en ligne quand les consommateurs sont ciblés, dans les téléphones mobiles et les assistants numériques pour la parole, la compréhension du langage et les tâches de traduction, dans les voitures sans conducteur pour la vision artificielle et la détection des obstacles, et dans le jeu de Go contre un champion que propose AlphaGo. L'apprentissage profond peut aussi être présent dans des applications moins connues, notamment dans la robotique et les prédictions de séismes. Une application comme TinEye (https://tineye.com) peut aussi vous être utile. Vous lui soumettez une image, et TinEye vous la retrouve sur l'Internet.

Commencer avec le miraculeux perceptron

L'algorithme de base du réseau neuronal est le neurone (aussi appelé unité). Des neurones en grand nombre organisés et reliés de manière à former une structure constituent un réseau neuronal, chaque neurone étant lié aux entrées et aux sorties d'autres neurones. Un neurone peut ainsi recevoir des données issues des exemples ou transmettre les résultats d'autres neurones, en fonction de sa localisation dans le réseau de neurones.

Frank Rosenblatt, du laboratoire d'aéronautique de Cornell, avait créé le premier exemple de neurone de ce type, le perceptron, il y a quelques décennies. C'était en 1957, sous le parrainage de l'United States Naval Research Laboratory (NRL). Rosenblatt était un

psychologue et un pionnier dans le domaine de l'intelligence artificielle. Il était expert en sciences cognitives, et son idée était de créer un ordinateur capable d'apprendre par tâtonnements, tout comme un être humain apprend.

Le perceptron n'était qu'un moyen subtil de tracer une ligne de séparation dans un espace élémentaire créé par les données d'entrée, comme le montre la Figure 11.1, sur laquelle sont représentées deux propriétés (en l'occurrence, la taille de l'animal et son niveau de domestication) à utiliser pour distinguer deux classes (dans cet exemple, les chats et les chiens). La formulation du perceptron produit une ligne dans un espace cartésien, et les exemples se répartissent plus ou moins parfaitement entre les groupes. L'approche est similaire à celle de Naive Bayes (voir Chapitre 10), dans laquelle la somme des probabilités conditionnelles est multipliée par les probabilités générales pour les besoins de la classification.

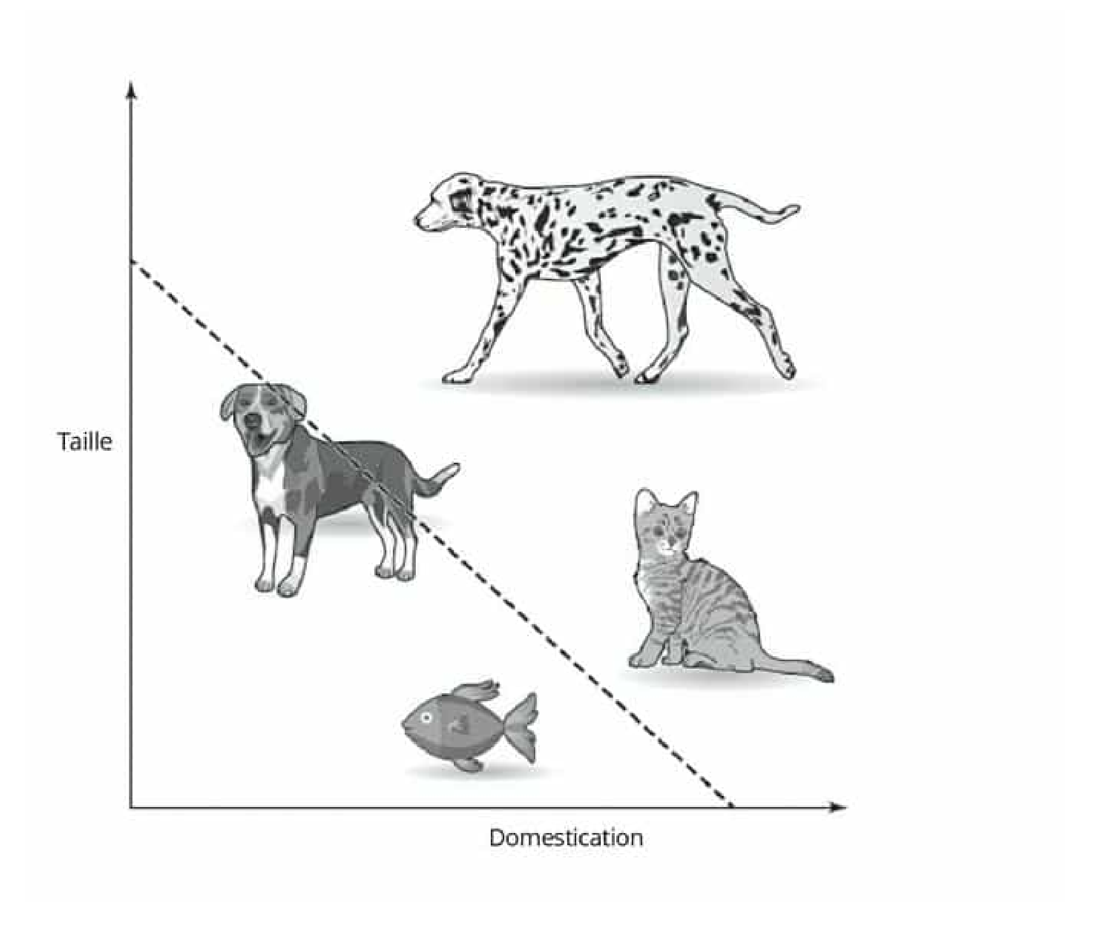


FIGURE 11.1 : Exemple de perceptron dans des tâches de classification simples et difficiles.

Le perceptron n'a pas répondu pleinement aux attentes de son créateur, ni à celles de ses bailleurs de fonds. Sa capacité est rapidement apparue limitée, même dans son domaine de spécialisation qu'était la reconnaissance d'image. La déception générale a débouché sur le premier hiver de l'IA et sur l'abandon du connexionnisme jusqu'aux années 1980. Des recherches ont cependant continué, bien que les crédits aient été coupés (Nils J. Nilsson, aujourd'hui à la retraite mais auparavant professeur d'IA à Stanford, donne plus de détails sur les progrès réalisés au cours de cette période dans l'article suivant : https://www.singularityweblog.com/ai-is-so-hot-weve-forgotten-all-about-the-ai-winter/).

Par la suite, des spécialistes ont tenté de créer un perceptron plus élaboré, et ils ont réussi. Les neurones dans un réseau neuronal représentent une évolution ultérieure du perceptron : ils sont nombreux, ils se connectent les uns aux autres, et ils imitent nos neurones quand ils s'activent sous l'effet d'un certain stimulus. En observant les fonctionnalités du cerveau humain, les scientifiques ont remarqué que les neurones recevaient des signaux mais n'en réémettaient pas toujours. La réémission d'un signal dépend de la quantité de signaux reçus. Quand un neurone reçoit suffisamment de stimuli, il émet une réponse ; dans le cas contraire, il reste inactif. De façon similaire, les neurones algorithmiques, une fois qu'ils ont reçu des données, les additionnent et les utilisent comme une fonction d'activation pour évaluer le résultat. Si l'input reçu atteint un certain seuil, le neurone transforme et transmet la valeur entrée ; sinon, il meurt tout simplement.



Pour émettre un résultat, les réseaux de neurones utilisent des fonctions spéciales appelées *fonctions d'activation*. Sachez simplement qu'il s'agit d'un composant essentiel des réseaux neuronaux, sans lequel ces derniers ne pourraient pas résoudre des problèmes complexes. Ces fonctions, telles des portes, laissent le signal passer, ou bien l'arrêtent. Cependant, elles ne se contentent pas de le laisser passer : elles le transforment de façon utile. L'apprentissage profond, par exemple, n'est pas possible en l'absence de fonctions d'activation efficaces comme l'unité linéaire rectifiée (*Rectified Linear Unit*, ou ReLU). Les fonctions d'activation constituent donc un aspect important du problème.

Imiter le cerveau qui apprend

Dans un réseau de neurones, il convient d'étudier tout d'abord l'architecture, c'est-à-dire la manière dont les

composants sont disposés. Les sections qui suivent sont consacrées aux considérations architecturales relatives aux réseaux neuronaux.

Étudier des réseaux neuronaux simples

Contrairement aux autres algorithmes, dans lesquels une séquence logique détermine la manière dont les données sont reçues et traitées, les réseaux neuronaux vous laissent décider vous-même comment l'information doit circuler en fixant le nombre d'unités (les neurones) et leur répartition en couches, ce que l'on appelle *l'architecture du réseau* (Figure 11.2).

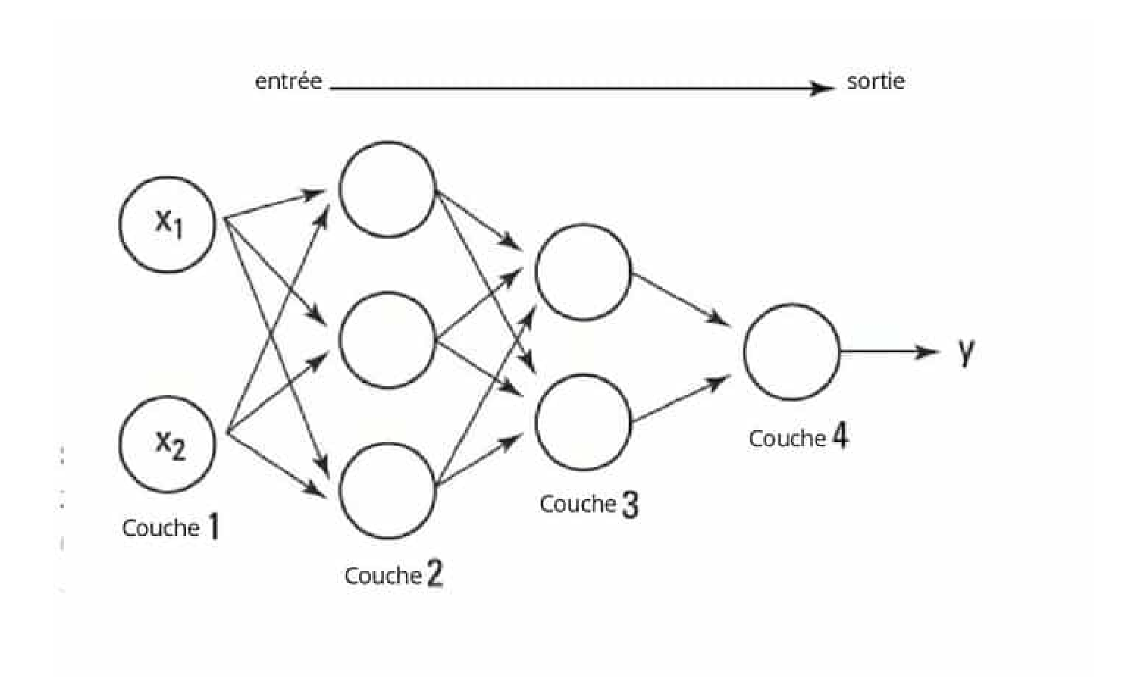


FIGURE 11.2 : Une architecture de réseau de neurones, de l'entrée à la sortie.

Cette figure représente l'architecture d'un réseau neuronal simple. Remarquez la manière dont les couches filtrent et

traitent l’information de façon progressive. Il s’agit d’une *circulation vers l’avant*, car les données circulent dans un seul sens. Les connexions lient exclusivement les unités d’une couche aux unités de la couche suivante (l’information circulant de gauche à droite). Il n’existe pas de connexion entre deux unités d’une même couche ni entre une unité et une autre unité en dehors de la couche suivante. Par ailleurs, l’information progresse dans un seul sens (de la gauche vers la droite). Les données traitées ne reviennent jamais vers les couches de neurones précédentes.

On utilise un réseau neuronal comme on utilise un système de filtrage stratifié de l’eau : de même qu’on verse l’eau au-dessus du système pour qu’elle descende et arrive filtrée dans la partie inférieure, sans qu’elle puisse remonter ni circuler latéralement, les éléments de données ne peuvent que circuler dans un sens à travers le réseau, en se mélangeant comme l’impose l’architecture. En utilisant la meilleure architecture pour mélanger les éléments, le réseau neuronal crée de nouveaux éléments composés au niveau de chaque couche et permet d’obtenir de meilleures prédictions. Il n’existe malheureusement aucun moyen de déterminer la meilleure architecture sans devoir tester empiriquement différentes solutions pour voir si les données sortantes permettent de prédire les valeurs cibles en sortie de réseau.



Les concepts peuvent parfois être mieux compris quand ils sont directement testés dans la réalité. Google propose un banc d’essai des réseaux neuronaux (http://playground.tensorflow.org) dans lequel vous pouvez tester de façon intuitive le fonctionnement d’un réseau de neurones. Pour ce faire, vous ajoutez ou supprimez des couches et vous modifiez les activations.

S’apercevoir que le secret est dans les poids

Les réseaux neuronaux présentent différentes couches, et chacune a ses propres poids. Les poids représentent la force de la connexion entre les neurones dans le réseau. Quand le poids de la connexion entre deux couches est réduit, cela signifie que le réseau déverse des valeurs qui circulent entre ces deux couches, et que suivre ce chemin ne devrait pas influencer la prédiction finale. De même, une forte valeur positive ou négative affecte les valeurs que reçoit la couche suivante, ce qui détermine certaines prédictions. Cette approche est analogue à celle des cellules du cerveau, qui ne sont pas isolées mais se connectent à d’autres cellules. Avec l’expérience, les connexions entre les neurones ont tendance à s’affaiblir ou à se renforcer pour activer ou désactiver certaines régions cellulaires du réseau cérébral, ce qui déclenche un autre traitement ou une activité (une réaction à un danger, par exemple, si l’information traitée signale une situation dans laquelle la vie est en jeu).

Chaque couche successive d’unités du réseau de neurones traite progressivement les valeurs provenant des éléments qui circulent, comme sur un tapis roulant. À mesure qu’elles sont transmises par le réseau, les données parviennent à chaque unité sous la forme d’une valeur totale calculée à partir des valeurs présentes dans la couche précédente, pondérées par les connexions vers la couche présente. Quand les données reçues des autres neurones dépassent un certain seuil, la fonction d’activation augmente la valeur stockée dans l’unité ; dans le cas contraire, elle met fin au signal en la réduisant. À l’issue du traitement par la fonction d’activation, le résultat est prêt à être transmis à travers la connexion vers la couche suivante. Ces étapes sont répétées pour chaque

couche, jusqu'à ce que les valeurs atteignent la fin du réseau, et l'on obtient alors le résultat.

Les poids des connexions et les fonctions d'activation constituent un moyen de mélanger et d'associer de façon créative les données entrées et traitées, et de créer ainsi de nouveaux éléments. Par ailleurs, l'activation rend non linéaire la recombinaison résultante des intrants reçus par les connexions. Ces deux types de composants du réseau neuronal permettent à l'algorithme d'assimiler des fonctions cibles complexes qui représentent la relation entre les éléments d'entrées et le résultat cible.

Comprendre le rôle de la rétropropagation

Dans le cerveau humain, l'apprentissage résulte de la formation et de la modification des synapses entre les neurones, à partir des stimuli reçus au cours d'expériences d'essais et d'erreurs. Les réseaux neuronaux sont un moyen de reproduire ce processus sous forme d'une formulation mathématique appelée la rétropropagation. Voici comment cette architecture composée d'unités de calcul interconnectées peut résoudre les problèmes : les unités reçoivent un exemple, et si elles ne prédisent pas correctement, elles reprennent le problème dans le système des poids existants en utilisant la rétropropagation et en y remédiant grâce à la modification de certaines valeurs. Ce processus progresse par itérations jusqu'à ce que le réseau neuronal puisse apprendre. Dans un réseau neuronal, les itérations s'appellent des époques, un nom approprié sachant qu'un réseau neuronal peut avoir besoin de plusieurs jours ou de plusieurs semaines d'apprentissage pour pouvoir apprendre des tâches complexes.



La rétropropagation fait appel à des calculs avancés qui nécessitent la maîtrise de notions comme les dérivées. La rétropropagation est un concept assez intuitif, sachant qu'elle ressemble à ce que l'on fait quand on exécute une tâche en tâtonnant et en procédant par itérations.

Depuis l'apparition de l'algorithme de rétropropagation dans les années 1970, les développeurs l'ont modifié à maintes reprises et il est actuellement envisagé de le repenser (vous pouvez lire l'avis de Geoffrey Hinton, un des coauteurs de la méthode, à l'adresse `https://www.axios.com/ai-pioneer-advocates-starting-over-2485537027.html`). La rétropropagation est au cœur de la renaissance actuelle de l'IA. Dans le passé, chaque progrès dans le processus d'apprentissage des réseaux neuronaux a débouché sur de nouvelles applications et a suscité un regain d'intérêt pour cette technique. Par ailleurs, la révolution actuelle de l'apprentissage profond, qui s'accompagne d'une renaissance des réseaux neuronaux (abandonnés au début des années 1990), est le résultat des principales avancées dans la manière dont les réseaux de neurones apprennent de leurs erreurs.

Introduire l'apprentissage profond

Après la rétropropagation, le progrès des réseaux neuronaux a amené l'apprentissage profond. Les recherches se sont poursuivies malgré l'hiver de l'IA, et les réseaux neuronaux ont surmonté des problèmes techniques comme la *disparition du gradient*, qui limite leur dimension. Pour pouvoir résoudre certains problèmes, les développeurs avaient besoin de réseaux neuronaux si étendus qu'ils étaient inenvisageables dans les années 1980. Par ailleurs, les chercheurs ont alors pu

profiter de l’évolution des CPU et des GPU (les processeurs graphiques, connus surtout pour leur application aux jeux vidéo).



La disparition du gradient est ce qui se produit quand le signal qu’on veut transmettre à travers un réseau de neurones se réduit rapidement à des valeurs proches de zéro : il ne passe plus par les fonctions d’activation. La raison à cela est que les réseaux neuronaux sont des enchaînements de multiplications, et chaque multiplication au-dessous de zéro fait diminuer rapidement les valeurs. Or, les fonctions d’activation ne laissent passer le signal que si les valeurs sont suffisamment élevées. Plus une couche de neurones est éloignée de la sortie, plus l’information risque d’être bloquée, les signaux étant trop faibles, si bien que les fonctions d’activation les stoppent. Par conséquent, le réseau cesse globalement d’apprendre, ou bien il apprend à un rythme considérablement ralenti.

De nouvelles solutions permettent d’éviter le problème de la disparition du gradient et un certain nombre d’autres problèmes techniques et d’avoir des réseaux plus vastes et plus profonds, ce qui fait contraste avec les réseaux superficiels plus simples du passé. Les réseaux profonds deviennent possibles grâce aux études menées par des scientifiques de l’Université de Toronto, au Canada, comme Geoffrey Hinton (https://www.utoronto.ca/news/artificial-intelligence-u-t), qui a mis l’accent sur le travail avec les réseaux neuronaux, même lorsqu’ils semblaient représenter une méthode d’apprentissage machine plutôt dépassée.

Les GPU sont des unités performantes de calcul matriciel et vectoriel qui sont nécessaires à la rétropropagation. Ces technologies rendent l’apprentissage des réseaux neuronaux viable à plus court terme et accessible à

davantage d’utilisateurs. Les recherches dans ce domaine ont aussi permis tout un ensemble de nouvelles applications. Les réseaux neuronaux peuvent apprendre à partir d’énormes volumes de données et tirer profit des « grandes données » (images, textes, transactions et données générées par les réseaux sociaux) en créant des modèles toujours plus performants, en fonction du flux de données qui les alimente.

Des acteurs importants comme Google, Facebook, Microsoft et IBM ont identifié cette nouvelle tendance et se sont intéressés aux nouveaux domaines de l’apprentissage profond à partir de 2012, en rachetant des sociétés et en engageant des spécialistes (Hinton travaille maintenant pour Google, et LeCun, le créateur des réseaux neuronaux convolutifs, dirige les recherches sur l’IA chez Facebook). Dans le cadre du projet Google Brain, dirigé par Andrew Ng et Jeff Dean, on a réuni 16000 ordinateurs pour calculer un réseau d’apprentissage profond comportant plus d’un milliard de poids, ce qui a débouché sur un apprentissage non supervisé à partir des vidéos de YouTube. Le réseau ainsi constitué a même pu déterminer de lui-même ce qu’est un chat, sans aucune intervention humaine (lire cet article de *Wired* : https://www.wired.com/2012/06/google-x-neural-network/).

COMPRENDRE LES PROBLÈMES D'APPRENTISSAGE PROFOND

Actuellement, les gens n'ont pas une vision réaliste de la façon dont l'apprentissage profond peut être utile à la société de façon globale. Quand une application de l'apprentissage profond bat un joueur aux échecs, on se demande quels autres prodiges cette technologie peut accomplir. Le problème est que même ses promoteurs ne comprennent pas très bien en quoi consiste l'apprentissage profond. Souvent, dans les articles techniques qui portent sur l'apprentissage profond, l'auteur décrit des couches de traitement nébuleux organisées en réseau, sans jamais évoquer ce qui se passe réellement en chacun de ces points. Ce qu'il est essentiel de retenir, c'est qu'en fait, l'apprentissage profond ne consiste pas à comprendre quoi que ce soit. Il consiste à utiliser un nombre considérable d'exemples pour en tirer des correspondances à partir de statistiques et de règles mathématiques. Quand une IA gagne une partie d'un jeu comportant un labyrinthe, elle ne comprend pas le concept de labyrinthe : elle sait simplement que certains intrants, manipulés d'une façon particulière, permettent d'aboutir à certains résultats gagnants.

Contrairement aux humains, l'apprentissage profond doit s'appuyer sur un nombre considérable d'exemples pour pouvoir découvrir des relations particulières entre les entrées et les sorties. Si l'on explique à un enfant que toute personne d'une certaine tranche d'âge n'est plus un enfant, pas encore un adolescent, mais un « préado », il sera capable d'identifier les individus relevant de cette catégorie avec un pourcentage d'exactitude élevé, même lorsqu'il s'agit d'un inconnu. Pour pouvoir accomplir la même tâche, une IA d'apprentissage profond nécessiterait une procédure d'apprentissage particulière et serait facilement induite en erreur, car les exemples qui seraient étrangers à son expérience ne compteraient pas.

Les humains peuvent aussi créer des hiérarchies du savoir même sans aucun apprentissage. Nous savons, par exemple, sans avoir besoin d'y réfléchir, que les chiens comme les chats sont des animaux. En outre, sachant que les chiens et les chats sont des animaux, une personne humaine peut facilement, par analogie, voir les autres animaux comme tels, même sans passer par un apprentissage spécifique. Dans l'apprentissage profond, il faudrait un apprentissage distinct pour chaque espèce ou genre animal. En résumé, l'apprentissage profond ne peut pas projeter ce qu'il sait sur d'autres situations, comme les humains savent le faire.

Même avec ces limites, l'apprentissage profond est un outil formidable, mais il ne doit pas être le seul outil de l'IA. L'utilisation de l'apprentissage profond pour discerner des structures ou des tendances là où les humains ne peuvent pas les voir est le parfait moyen d'appliquer cette technologie. Les structures et les tendances sont un élément essentiel de la découverte de nouveaux faits. Par exemple, tester des mélanges pour lutter contre le cancer pourrait demander un temps considérable. En voyant des tendances là où les humains ne les voient pas, l'apprentissage profond pourrait permettre de réaliser des progrès décisifs vers une solution, avec bien moins d'efforts que ce qui serait nécessaire autrement.

Expliquer la différence dans l'apprentissage profond

L'apprentissage profond peut sembler n'être qu'un réseau neuronal plus vaste, mobilisant un plus grand nombre d'ordinateurs : en d'autres termes, une simple percée dans le domaine des mathématiques et de la puissance de traitement informatique permettant de disposer de réseaux d'aussi grande dimension. Or, dans

l’apprentissage profond, un changement de nature qualitative s’est produit par rapport aux réseaux neuronaux superficiels. C’est davantage que le changement de paradigme de brillants ingénieurs. L’apprentissage profond change le paradigme de l’apprentissage machine, avec le passage de la création d’éléments (des éléments qui facilitent l’apprentissage et qu’il faut créer en recourant à l’analyse de données) à l’apprentissage d’éléments (éléments complexes créés automatiquement en fonction des éléments réels). Cet aspect n’était pas perçu en utilisant des réseaux plus petits, mais il devient évident quand on utilise un réseau neuronal constitué de nombreuses couches avec une grande quantité de données.

En examinant de plus près ce qu’est l’apprentissage profond, on peut être surpris d’y trouver de la technologie ancienne en masse, mais curieusement, tout fonctionne comme jamais auparavant. Les chercheurs ayant fini par trouver comment faire fonctionner ensemble de bonnes vieilles solutions toutes simples, le système des grandes données peut automatiquement filtrer, traiter et transformer les données. Ainsi, des nouvelles activations comme ReLU ne sont pas si nouvelles : elles datent de l’époque du perceptron. Les capacités de reconnaissance d’image qui, au début, ont tellement contribué à la popularité de l’apprentissage profond, ne sont pas nouvelles non plus. À l’origine, l’apprentissage profond a connu son essor grâce aux réseaux neuronaux convolutifs. Découverts dans les années 1980 par le scientifique français Yann LeCun (dont la page d’accueil personnelle est http://yann.lecun.com/), ces réseaux donnent aujourd’hui des résultats étonnants, parce qu’ils utilisent un grand nombre de couches neuronales et de grandes quantités de données. Il en est de même de la technologie qui permet à une machine de comprendre les paroles des personnes ou de traduire un texte d’une langue à une autre : il s’agit d’une technologie

datant de plusieurs décennies, qu'un chercheur a revisitée et adaptée au nouveau paradigme de l'apprentissage profond.

Naturellement, la différence provient aussi en partie des données (nous y reviendrons), de l'utilisation accrue des GPU et du recours aux réseaux informatiques. Conjointement avec le parallélisme (davantage d'ordinateurs organisés en grappes et fonctionnant en parallèle), les GPU permettent de créer des réseaux plus vastes et de les faire travailler sur davantage de données. On estime qu'un GPU peut exécuter certaines opérations 70 fois plus vite que n'importe quel CPU, ce qui permet de réduire les temps d'apprentissage des réseaux neuronaux à quelques jours, voire quelques heures, au lieu de plusieurs semaines.



Pour plus d'informations sur la dynamisation de l'apprentissage machine grâce au GPU dans le cadre de l'utilisation d'un réseau de neurones, consultez ce document technique : https://icml.cc/2009/papers/218.pdf.

Trouver des solutions encore plus subtiles

L'apprentissage profond influence l'efficacité de l'IA pour résoudre des problèmes dans la reconnaissance d'images, la traduction automatique et la reconnaissance vocale qui étaient initialement résolus par l'IA classique et l'apprentissage machine. En outre, il apporte des solutions nouvelles et avantageuses :

- la continuité grâce à l'apprentissage incrémentiel ;
- des solutions réutilisables grâce à l'apprentissage par transfert ;

» une plus grande démocratisation de l'IA grâce à des cadres de travail en source ouverte ;

» des solutions simples et directes grâce à l'*apprentissage de bout en bout*.

Utiliser l'apprentissage incrémentiel

Les réseaux neuronaux sont plus flexibles que d'autres algorithmes d'apprentissage machine, et ils peuvent continuer à apprendre en produisant des prédictions et des classifications. Cette capacité provient des algorithmes d'optimisation qui permettent aux réseaux neuronaux d'apprendre, sachant qu'ils peuvent travailler de façon répétitive sur de petits échantillons d'exemples (*apprentissage par lots*) ou même sur des exemples simples (*apprentissage incrémentiel*). Les réseaux d'apprentissage profond peuvent bâtir leur connaissance pas à pas en étant réceptifs aux nouvelles informations qui peuvent arriver (comme un bébé est toujours réceptif aux nouveaux stimuli et aux nouvelles expériences d'apprentissage). Une application de l'apprentissage profond sur un site de réseaux sociaux, par exemple, peut apprendre à reconnaître les chats sur les images. À mesure que des photos de chats sont publiées par les utilisateurs, l'application les reconnaît et leur affecte l'étiquette appropriée. Quand les gens se mettent à publier des photos de chiens, le réseau de neurones n'a pas besoin d'apprendre à nouveau : il peut continuer en apprenant de la même manière à reconnaître les photos de chiens. Cette capacité est particulièrement utile pour faire face à la variabilité des données de l'Internet. Un réseau d'apprentissage profond peut être ouvert à la nouveauté et adapter ses poids en conséquence.

Utiliser l'apprentissage par transfert

La flexibilité est pratique même quand un réseau termine son apprentissage, mais il faut le réutiliser dans une optique différente de l'apprentissage initial. Les réseaux qui distinguent les objets et les classent correctement ont besoin de beaucoup de temps et d'une puissance de calcul considérable pour apprendre quoi faire. Quand la capacité d'un réseau est étendue à de nouvelles sortes d'images qui ne faisaient pas partie de l'apprentissage précédent, la connaissance doit être transférée vers ce nouveau problème (apprentissage par transfert).

Ainsi, on peut transférer un réseau qui est capable de faire la distinction entre les chiens et les chats pour qu'il exécute une tâche dans le cadre de laquelle il devra distinguer des assiettes de macaronis au fromage. La majorité des couches du réseau sont utilisées telles quelles (elles sont figées), puis on travaille sur les couches finales, celles de la sortie (peaufinage). En un court laps de temps et avec un plus petit nombre d'exemples, le réseau appliquera aux macaronis au fromage ce qu'il a appris à faire avec les chiens et les chats. Il fera même mieux qu'un réseau neuronal ayant seulement appris à reconnaître les macaronis et le fromage.

L'apprentissage par transfert est une chose nouvelle pour la plupart des algorithmes d'apprentissage machine, et il ouvre un marché possible pour transférer la connaissance d'une application vers une autre, d'une entreprise vers une autre. C'est ce que fait déjà Google, qui partage son immense entrepôt de données en rendant publics les réseaux qu'il y a créés (plus de détails dans cette publication : https://techcrunch.com/2017/06/16/objectdetection-api/). Il s'agit d'une étape dans la démocratisation de l'apprentissage profond, consistant à permettre à tout le monde d'accéder à ses potentialités.

Une démocratisation par le recours à des cadres en source ouverte

Aujourd'hui, les réseaux peuvent être accessibles à tout le monde, avec l'accès à des outils de création de réseaux d'apprentissage profond. Il ne s'agit pas simplement de mettre sur la place publique des articles scientifiques qui expliquent comment fonctionne l'apprentissage profond : il s'agit de programmation. Dans les tout débuts de l'apprentissage profond, il fallait créer chaque réseau à partir de rien, sous forme d'une application développée dans un langage comme le C++, ce qui en limitait l'accès à un petit nombre de spécialistes. Les capacités actuelles en matière de création de scripts (par exemple avec Python : voir http://www.python.org) constituent un progrès, en raison du vaste éventail de cadres d'apprentissage profond en source ouverte comme TensorFlow, de Google (https://www.tensorflow.org/) ou PyTorch, de Facebook (http://pytorch.org/). Ces cadres permettent la réplication des plus récentes avancées dans l'apprentissage profond à l'aide de commandes simples.



Toute médaille a ses revers. Les réseaux neuronaux ont besoin d'énormes volumes de données pour pouvoir fonctionner, or les données ne sont pas accessibles à tout le monde car elles sont détenues par de grandes organisations. L'apprentissage par transfert peut pallier le manque de données, mais de façon partielle seulement, car certaines applications ne peuvent pas fonctionner sans les données réelles. En conséquence, la démocratisation de l'IA a ses limites. Par ailleurs, les systèmes d'apprentissage profond sont si complexes que leurs résultats sont à la fois difficiles à expliquer (ce qui laisse la porte ouverte aux biais et à la discrimination) et fragiles, car il est possible de trouver des trucs pour en fausser le fonctionnement (voir

https://www.dvhardware.net/article67588.html). Tout réseau de neurones peut être sensible aux attaques, ou à des manipulations en entrée visant à tromper le système pour qu'il donne une mauvaise réponse.

Utiliser l'apprentissage de bout en bout

Enfin, l'apprentissage profond permet l'*apprentissage de bout en bout*, ce qui signifie qu'il résout des problèmes de façon plus facile et plus simple que la solution précédente d'apprentissage profond. Cette façon de résoudre les problèmes peut avoir un plus grand impact, qu'il s'agisse de reconnaître des visages connus ou de conduire une voiture. Avec l'IA classique, il fallait diviser le problème en sous-problèmes plus faciles à gérer afin d'obtenir un résultat acceptable dans un délai viable. S'il s'agissait de reconnaître des visages sur une photo, par exemple, les systèmes d'IA précédents décomposaient le problème comme suit :

1. **Trouver les visages sur la photo.**
2. **Recadrer les visages.**
3. **Traiter les visages recadrés de manière à obtenir une pose similaire à celle d'une photo d'identité.**
4. **Alimenter le réseau neuronal avec les visages recadrés et traités, de manière à en faire autant d'exemples pour l'apprentissage de la reconnaissance d'image.**

Aujourd'hui, on peut soumettre la photo à une architecture d'apprentissage profond et guider celle-ci pour qu'elle apprenne à trouver les visages sur les images, puis à les classer. On peut utiliser la même approche pour la traduction automatique, la reconnaissance vocale, et

même les voitures sans conducteur (voir Chapitre 14). Dans tous les cas, on transmet simplement l'information d'entrée à un système d'apprentissage profond et l'on obtient le résultat voulu.

Détecter les contours et les formes sur les images

Les réseaux neuronaux convolutifs (aussi appelés en anglais ConvNet ou CNN) ont contribué à la récente renaissance de l'apprentissage profond. Les sections qui suivent expliquent comment ils permettent de détecter les contours et les formes sur les images, dans le cadre de tâches comme le déchiffrage de textes manuscrits.

Commencer par la reconnaissance de caractères

Les réseaux neuronaux convolutifs ne sont pas une idée nouvelle. Ils sont apparus à la fin des années 1980 et sont dus à Yann LeCun (aujourd'hui directeur de l'IA chez Facebook), qui travaillait alors dans la division recherche d'AT & T avec Yoshua Bengio, Leon Bottou et Patrick Haffner sur un réseau appelé LeNet5. Vous pouvez voir ce réseau à l'adresse http://yann.lecun.com/exdb/lenet/ ou sur cette vidéo, dans laquelle LeCun lui-même, bien plus jeune qu'aujourd'hui, fait une démonstration : https://www.youtube.com/watch?v=FwFduRA_L6Q. À l'époque, disposer d'une machine capable de reconnaître des chiffres écrits à la main était une chose vraiment exceptionnelle. Cela a permis la détection automatique du code postal pour le tri du

courrier et le tri des courriers électroniques entrants et sortants.

Des développeurs avaient obtenu des résultats en associant un réseau neuronal à des images de chiffres. À chaque pixel de l'image correspondait un nœud dans le réseau. Le problème avec cette méthode est que le réseau ne peut pas assurer l'*invariance par translation*, c'est-à-dire la capacité de déchiffrer le nombre en cas de variation de taille, de distorsion ou de position différente sur l'image, comme le montre la Figure 11.3. Un réseau neuronal similaire ne pouvait détecter que des nombres similaires : ceux qu'il avait déjà vus auparavant. Il commettait aussi de nombreuses erreurs. Transformer l'image avant de la soumettre à un réseau neuronal permettait de résoudre le problème en partie : il s'agissait de redimensionner, de déplacer et d'épurer les pixels, et de créer des blocs d'information particuliers pour que le traitement par le réseau soit plus performant. Cette technique, appelée *création d'éléments*, nécessite à la fois des compétences par rapport aux transformations d'images nécessaires et de nombreux calculs en termes d'analyse de données. À l'époque, les tâches de reconnaissance d'image étaient plus de l'artisanat qu'un travail de scientifique.

Les convolutions ont facilement résolu le problème de l'invariance par translation, sachant qu'elles constituent une approche différente du traitement d'image à l'intérieur du réseau neuronal. Les convolutions sont le fondement de LeNet5 et elles constituent les éléments de base pour tous les vrais réseaux neuronaux convolutifs qui réalisent les tâches suivantes :

- **Classification d'images :** déterminer quel objet apparaît sur une image.
- **Détection d'images :** localiser un objet sur une image.

» **Segmentation d’images :** séparer les différentes parties d’une image en fonction de leur contenu ; par exemple, si l’image représente une route, séparer la route elle-même des voitures et des piétons.

Expliquer comment fonctionnent les convolutions

Pour comprendre comment fonctionnent les convolutions, il faut partir de l’intrant, qui est une image composée d’une ou plusieurs couches de pixels appelées les *canaux*, avec des valeurs de pixels comprises entre 0 (pixel complètement activé) et 256 (pixel éteint). Les images codées selon le format RVB, par exemple, utilisent un canal pour le rouge, un autre pour le vert et un troisième pour le bleu. Le mélange de ces canaux donne la palette des couleurs que l’on voit sur l’écran.

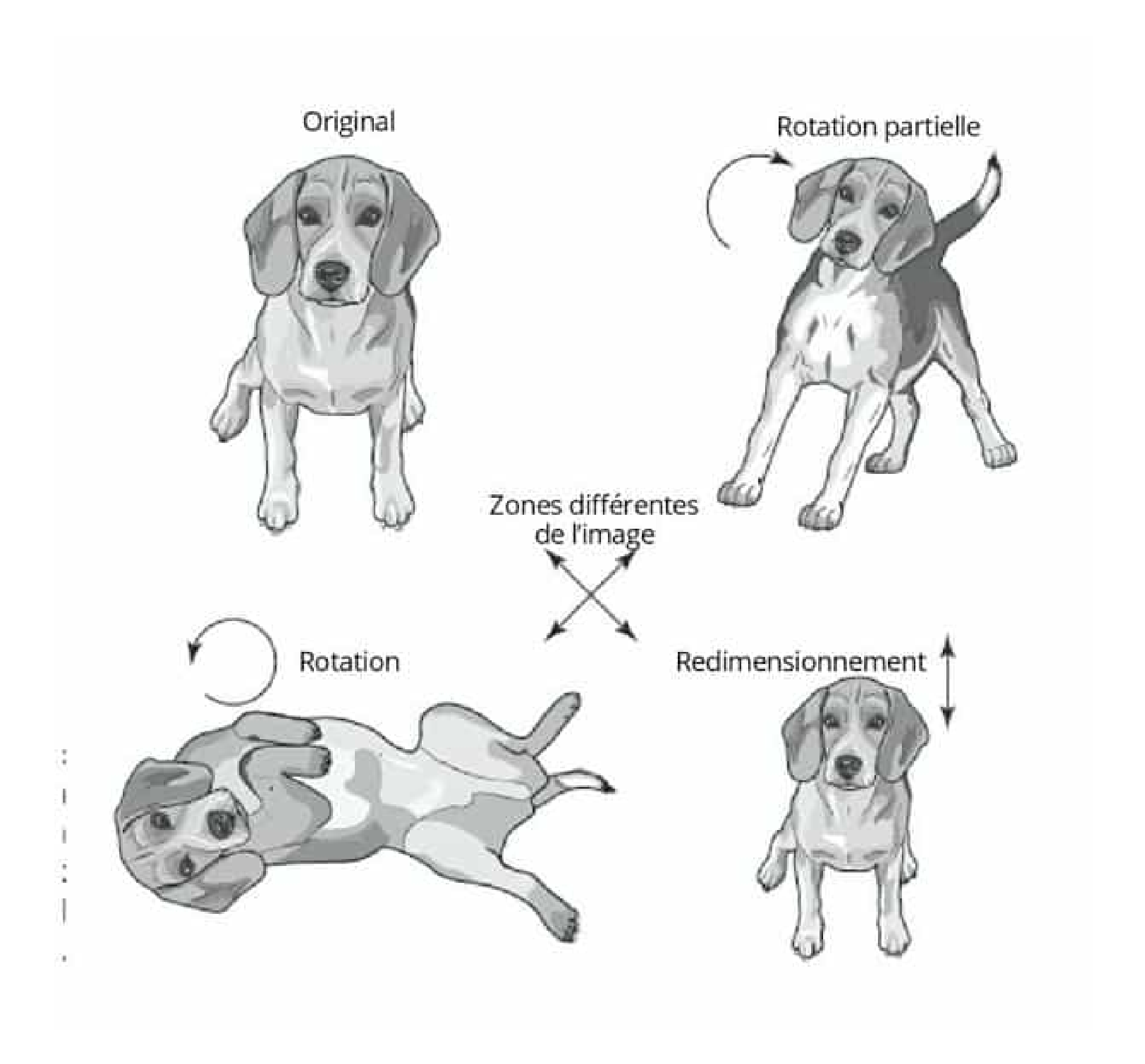


FIGURE 11.3 : Avec l'invariance par translation, un réseau neuronal identifie le chien et ses variations.

Les données d'entrée subissent des transformations simples pour redimensionner les valeurs des pixels (par exemple, pour redéfinir une échelle de 0 à 1), puis les valeurs sont traitées. Ces transformations des données facilitent le travail des convolutions, sachant que les convolutions sont simplement des opérations de multiplication et d'addition (Figure 11.4). La couche neuronale de la convolution prend des petites parties de l'image, multiplie les valeurs des pixels dans la partie traitée par une grille de nombres spécifiquement définis, additionne tous les résultats de la multiplication, et projette le résultat sur la prochaine couche neuronale.

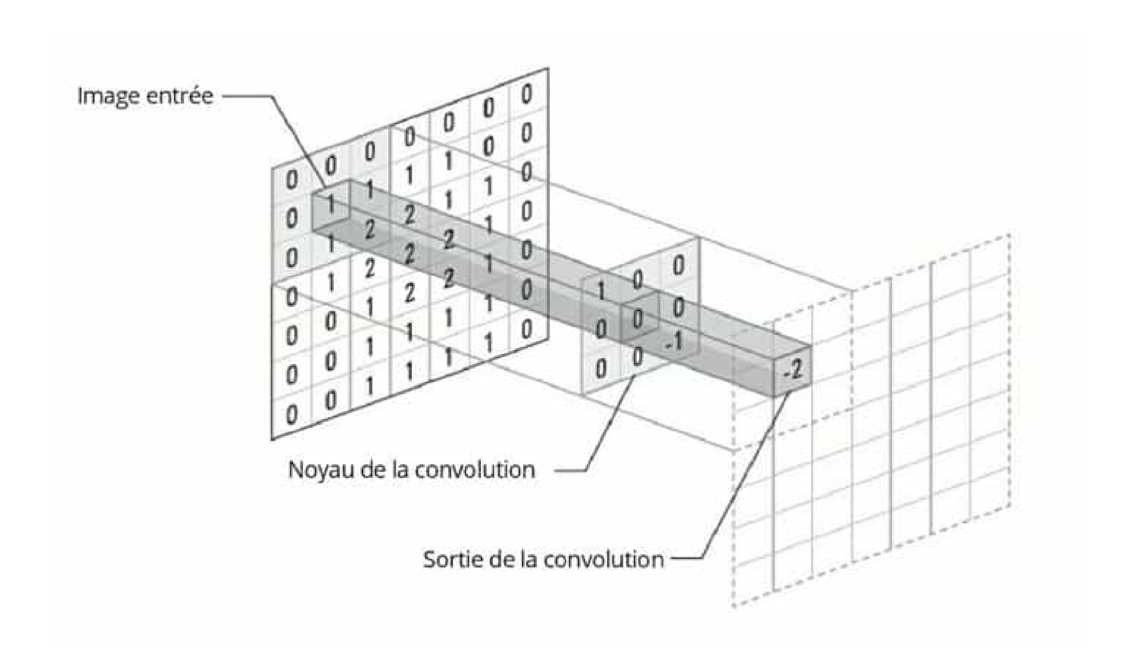


FIGURE 11.4 : Une convolution qui balaye une image.

Une telle opération est flexible, car la rétropropagation constitue la base de la multiplication à l'intérieur de la convolution (pour des précisions sur l'étape de la convolution, avec une animation, voir l'article à l'adresse `https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/`), et les valeurs que la convolution filtre sont les caractéristiques de l'image, qui sont importantes pour permettre au réseau neuronal de réussir sa tâche de classification. Certaines convolutions ne capturent que les lignes, d'autres seulement les courbes ou des structures particulières, quel que soit l'endroit où elles apparaissent sur l'image (et c'est la propriété d'invariance par translation des convolutions). À mesure que les données de l'image traversent différentes convolutions, elles sont transformées, rassemblées et forment des structures incroyablement complexes, jusqu'à ce que la convolution produise des images de référence (par exemple, l'image d'un chat ou d'un chien typique), que le réseau neuronal convolutif va ensuite utiliser pour détecter de nouvelles images.



Pour en savoir plus sur les convolutions, intéressez-vous à cette visualisation créée par des chercheurs de chez Google dans le cadre des projets Research et Google Brain. La visualisation montre le fonctionnement interne de GoogleLeNet, un réseau de 22 couches développé par des chercheurs chez Google (voir la page https://distill.pub/2017/feature-visualization/). Dans l'annexe (https://distill.pub/2017/feature-visualization/appendix/), les auteurs présentent des exemples à partir des couches conçues pour détecter d'abord les contours, puis les textures, puis les structures entières, puis les parties, et enfin, les objets entiers.

Singulièrement, la mise en place des architectures de base de ConvNet n'est pas difficile. Dites-vous simplement qu'il est préférable d'avoir le plus grand nombre possible de couches. Il s'agit de déterminer le nombre de couches de convolution et certaines caractéristiques comportementales de la convolution, comme la constitution de la grille (filtre, noyau ou détecteur d'éléments), la manière dont la grille traite l'image (le pas, ou *stride*), et la manière dont l'algorithme traite les bords de l'image (la marge, ou *padding*).



Quand on se penche sur le fonctionnement des convolutions, on constate que dans l'apprentissage profond, les données subissent des transformations plus profondes qu'avec n'importe quel algorithme d'apprentissage machine ou avec un réseau neuronal superficiel. Plus les couches sont nombreuses, plus l'image subit de transformations, et plus il y a de profondeur.

Progresser par les défis de l'image

Les réseaux neuronaux convolutifs sont une invention brillante. AT & T a déjà utilisé LeNet5 pour les lecteurs de chèques des terminaux bancaires. Néanmoins, il y a eu un nouvel hiver de l'IA au milieu des années 1990, quand de nombreux chercheurs et investisseurs ont cessé de penser que les réseaux neuronaux allaient révolutionner l'IA. En outre, à cette époque le degré de complexité des données était insuffisant. Les chercheurs ont pu obtenir des résultats comparables à ceux de LeNet5 en utilisant de nouveaux algorithmes d'apprentissage machine appelés les machines à vecteurs de support (mis au point par la tribu des analogistes) et Random Forests, un développement fondé sur les arbres de décision dû à la tribu des symbolistes (voir Chapitre 10).

Seuls de rares chercheurs comme Geoffrey Hinton, Yann LeCun et Yoshua Bengio ont continué à développer des technologies de réseaux neuronaux jusqu'à ce qu'un nouvel ensemble de données permette une nouvelle percée et mette fin à l'hiver de l'IA. En parallèle, en 2006 la chercheuse Fei-Fei Li, qui enseigne l'informatique à l'Université de l'Illinois Urbana-Champaign (et qui est aujourd'hui responsable scientifique chez Google Cloud et professeur à Stanford) s'est employée à rassembler davantage de jeux de données réelles afin de pouvoir mieux tester les algorithmes. Elle a collecté un nombre considérable d'images de manière à disposer d'un vaste nombre de classes d'objets. Avec son équipe, elle est venue à bout de cette tâche colossale grâce à l'utilisation de Mechanical Turk, un service fourni par Amazon qui vous permet de demander à des gens d'effectuer pour vous des microtâches (comme classer une image) à peu de frais.

Le jeu de données résultant, réalisé en 2009, a reçu le nom d'ImageNet. Il comportait 3,2 millions d'images étiquetées, classées selon 5 247 catégories organisées de façon hiérarchique. Vous pouvez le découvrir sur la page

http://www.image-net.org/ ou lire l'article original qui présente le jeu de données sur la page http://www.image-net.org/papers/imagenet_cvpr09.pdf. ImageNet n'a pas tardé à faire son apparition dans un concours en 2010, à l'occasion duquel les réseaux neuronaux ont fait la preuve de leur capacité de classer correctement des images appartenant à un millier de classes différentes.

Ce concours a duré sept ans (il s'est terminé en 2017), et les algorithmes gagnants ont représenté un net progrès dans l'exactitude avec 97,3 % de prédiction correcte des images contre 71,8 % auparavant, une performance qui surpasse les capacités humaines (en effet, les humains commettent des erreurs quand ils classent des objets). Au début, les chercheurs ont remarqué que leurs algorithmes devenaient plus performants quand ils recevaient davantage de données (il n'existait encore rien qui ressemble à ImageNet), et ils ont commencé à tester de nouvelles idées et à améliorer les architectures de réseaux neuronaux.

Même si les concours ImageNet n'ont plus lieu, les chercheurs développent d'autres architectures de réseaux neuronaux convolutifs en améliorant la précision ou les capacités de détection ainsi que la robustesse du processus et des résultats. En fait, bien souvent, les solutions à base d'apprentissage profond sont au stade expérimental et ne sont pas encore adaptées à des applications critiques comme les activités bancaires ou la sécurité, non seulement à cause des difficultés que pose leur interprétation, mais aussi en raison des vulnérabilités possibles.



Les vulnérabilités prennent des formes variées. Les chercheurs ont constaté qu'en ajoutant un bruit spécialement étudié ou en modifiant un seul pixel dans une image, un réseau neuronal convolutif pouvait changer

radicalement ses réponses, dans des attaques non ciblées (il suffit de tromper le CNN) ou ciblées (en faisant en sorte que le CNN donne une réponse spécifique). Pour approfondir ce sujet, consultez le tutoriel OpenAI sur la page https://blog.openai.com/adversarial-example-research/. OpenAI est un groupe d'études à but non lucratif spécialisé dans l'IA. La lecture de l'article intitulé « One pixel attack for fooling deep neural networks » (https://arxiv.org/abs/1710.08864) est utile également. Le problème est que la technologie des réseaux neuronaux convolutifs n'est pas encore sécurisée. On ne peut pas se reposer sur elle, il convient de ne l'utiliser qu'avec une grande prudence.

Apprendre à imiter l'art et la vie

Les CNN ne concernent pas seulement la vision par ordinateur, ils sont importants également pour un certain nombre d'autres applications (ils sont nécessaires, par exemple, à la vision artificielle dans les voitures sans conducteur). Ils ont incité de nombreux chercheurs à investir du temps et des efforts dans la révolution de l'apprentissage profond. C'est ainsi que la recherche et le développement ont fait émerger de nouvelles idées. Les tests subséquents ont finalement permis une innovation dans le domaine de l'IA, les ordinateurs pouvant désormais comprendre le langage parlé, traduire des textes écrits dans des langues étrangères et créer des textes et des images modifiées, ce qui montre de quelle manière des calculs complexes effectués sur des distributions statistiques peuvent déboucher sur de l'art, de la créativité et de l'imagination. Quand on parle de l'apprentissage profond et de ses applications possibles, il convient de mentionner aussi les réseaux de neurones récurrents et les réseaux antagonistes génératifs, sans

quoi on ne saurait avoir une idée claire de ce que l'apprentissage profond peut apporter à l'IA.

Mémoriser des séquences qui comptent

Une des faiblesses que présentent les CNN est l'absence de mémoire. Un CNN est capable de comprendre une image donnée, mais lorsqu'il s'agit de comprendre une image dans un contexte, par exemple la trame d'une vidéo, les problèmes deviennent difficiles à surmonter pour l'IA et le système ne sait pas donner la bonne réponse. Les problèmes importants sont souvent des problèmes de séquences. Pour comprendre ce que raconte un livre, on le lit page par page. Le livre est une succession de séquences imbriquées. La page est une succession de mots, et chaque mot est une succession de lettres. Il s'agit de comprendre la séquence de lettres, de mots et de pages. Le type de système approprié est le réseau de neurones récurrents, qui traite les entrées réelles tout en se référant aux entrées passées. Les intrants ne font pas que transiter dans ce réseau comme dans n'importe quel autre réseau neuronal, ils effectuent aussi des boucles à l'intérieur du réseau. Tout se passe comme si le réseau percevait un écho de lui-même.

Quand on alimente un réseau de neurones récurrents avec une séquence de mots, il apprend à déterminer la manière de terminer une proposition selon la façon dont un mot est précédé par certains autres mots. Le réseau de neurones récurrents n'est pas simplement une technologie capable d'automatiser la compilation des intrants (comme lorsque le navigateur Web complète automatiquement la séquence de mots-clés que vous saisissez dans la requête). En outre, les réseaux de neurones récurrents peuvent traiter des séquences et produire en sortie une

traduction, comme le sens général d'une expression (une IA peut maintenant lever l'ambiguïté d'une expression là où la formulation est importante) ou la traduction d'un texte dans une autre langue (là encore, la traduction se fait dans un certain contexte). Cela fonctionne même avec les sons, car il est possible d'interpréter certaines modulations comme des mots. Les réseaux de neurones récurrents permettent aux ordinateurs et aux téléphones mobiles de comprendre, avec une grande précision, non seulement ce que vous avez dit (c'est avec la même technologie qu'on obtient un sous-titrage automatique), mais aussi ce que vous vouliez dire, ce qui ouvre la voie à des programmes qui discutent avec vous et à des assistants numériques comme Siri, Cortana et Alexa.

Découvrir la magie des conversations avec l'IA

Un *agent conversationnel* est un logiciel capable de converser avec vous grâce à deux méthodes : oralement (vous lui parlez et vous écoutez sa réponse) ou par écrit (vous tapez au clavier ce que vous voulez dire et vous lisez sa réponse). Vous en avez peut-être entendu parler sous un autre nom (chatbot, chatterbot, talkbot, *etc.*), mais peut-être en utilisez-vous déjà un sur votre téléphone intelligent, sur votre ordinateur ou sur un appareil spécial comme Siri, Cortana ou Alexa. Peut-être dialoguez-vous avec un agent conversationnel quand vous contactez le service clients d'une société sur Internet ou au téléphone, ou par l'intermédiaire d'une application sur votre téléphone mobile, ou quand vous utilisez Twitter, Slack, Skype ou d'autres applications dédiées à la conversation.

Les agents conversationnels représentent un enjeu commercial important car ils permettent aux entreprises de réaliser des économies tout en maintenant un contact

permanent avec les clients qu'elles servent – mais l'idée n'est pas nouvelle. Même si leur appellation est récente (le mot *chatbot* a été choisi en 1994 par Michael Mauldin, l'inventeur du moteur de recherche Lycos), ils sont considérés comme ce que l'IA fait de mieux. D'après Alan Turing, il ne devrait pas être possible de détecter une IA forte en parlant avec elle. Turing avait conçu un fameux test conversationnel pour déterminer si une IA avait acquis ou non une intelligence équivalente à celle d'un être humain.



Une IA est faible quand elle montre un comportement intelligent sans avoir une conscience analogue à celle d'un être humain. Une IA est forte quand elle peut réellement penser comme un être humain.

Le test de Turing consiste à faire communiquer une personne humaine avec deux sujets, l'un étant une personne et l'autre une machine, par l'intermédiaire d'un terminal. La conversation doit permettre à la personne de déterminer lequel des deux sujets est une IA. Turing affirmait que si une IA peut induire en erreur l'utilisateur en lui faisant croire qu'il converse avec une personne humaine, alors il est possible de penser que l'IA est comparable à une intelligence humaine. Il s'agit d'un problème difficile, car l'IA ne doit pas seulement répondre correctement et en utilisant une syntaxe grammaticale correcte, elle doit aussi intégrer le contexte (le lieu, le moment et les caractéristiques de la personne avec qui elle converse) et manifester une personnalité cohérente (elle doit apparaître comme une personne réelle, en termes de connaissances et d'attitude).

Depuis les années 1960, mettre en défaut le test de Turing s'est révélé être une motivation pour mettre au point des agents conversationnels, en se fondant sur le concept des modèles d'extraction d'information. Plus précisément, le traitement automatique du langage naturel consiste à traiter des éléments de langage entrés par l'utilisateur.

Certains mots ou groupes de mots appellent des réponses présélectionnées, et le système puise dans sa mémoire pour fournir les réponses.



Le *traitement automatique du langage naturel* est une analyse de données centrée sur le texte. L'algorithme découpe le texte en unités lexicales (noms, verbes et adjectifs) en éliminant toute information moins utile ou déroutante. Le texte découpé en unités lexicales est traité au moyen d'opérations statistiques ou d'un apprentissage machine. Ainsi, le traitement automatique du langage naturel permet d'identifier des mots et leur signification et de déterminer si un texte est similaire à un autre.

C'est en 1966 que Joseph Weizenbaum a créé le premier agent conversationnel de ce type, ELIZA, sous forme d'un psychothérapeute virtuel. ELIZA était constitué d'heuristiques simples, c'est-à-dire d'expressions de base à adapter au contexte, et de mots-clés qui appelaient une réponse appropriée parmi un ensemble déterminé de réponses possibles. Vous pouvez essayer en ligne une version d'ELIZA, à l'adresse http://www.masswerk.at/elizabot/. Vous serez peut-être surpris de lire cette conversation sérieuse qu'ELIZA a eue avec son créateur : http://www.masswerk.at/elizabot/eliza_test.html.

Ce type de modèle, basé sur la recherche de la bonne réponse à partie d'un ensemble d'associations (*retrieval-based model*), fonctionne bien quand on l'interroge sur des sujets présélectionnés, car il intègre une connaissance humaine, tout comme un système expert (voir Chapitre 3), si bien qu'il est capable de répondre en formulant une phrase pertinente et correcte d'un point de vue grammatical. Des problèmes surgissent quand les questions portent sur d'autres sujets. L'agent conversationnel peut essayer de parer la difficulté en

retournant la question à l'utilisateur sous une autre forme (comme le faisait ELIZA), et l'on se rend compte que l'on a affaire à un interlocuteur artificiel. Une solution consiste à créer de nouvelles phrases, par exemple par référence à des modèles statistiques, à l'apprentissage machine, voire à un réseau de neurones récurrents ayant déjà reçu un apprentissage *ad hoc*, et capable de générer un discours neutre ou même de refléter la personnalité de quelqu'un en particulier. On parle alors de *modèles* génératifs. Nous sommes aujourd'hui aux confins de la robotique, sachant que générer du langage à la volée n'est pas facile.

Les modèles génératifs ne répondent pas toujours par des phrases correctes et pertinentes, mais un certain nombre de chercheurs ont récemment réalisé des percées, surtout dans le domaine des réseaux de neurones récurrents. Comme mentionné précédemment, le secret réside dans la séquence : on fournit une séquence d'entrée dans un langage et une séquence de sortie dans un autre langage, comme dans un problème de traduction automatique. En l'occurrence, la séquence d'entrée et la séquence de sortie sont dans la même langue. La séquence d'entrée est un élément de la conversation, et la séquence de sortie est la réaction qui suit.

Compte tenu de l'état actuel de la technique en matière de développement d'agents conversationnels, les réseaux de neurones récurrents sont très performants pour les échanges courts, mais il est difficile d'obtenir des résultats parfaits avec des échanges plus longs ou avec des expressions plus compliquées. À l'instar des modèles de type *retrieval-based*, les réseaux de neurones récurrents réutilisent les informations qu'ils acquièrent, mais pas de manière organisée. Quand le champ du discours est limité, ces systèmes produisent de bonnes réponses, mais les choses se gâtent quand le contexte est ouvert et général, car il faudrait alors que le système dispose de connaissances comparables à celles qu'un être humain

acquiert toute sa vie durant (les humains savent bien converser en s’appuyant sur leur expérience et sur leurs connaissances).

Tout est dans les données avec lesquelles on alimente le réseau de neurones récurrents. Ainsi, Google Smart Reply, un agent conversationnel de Google, donne des réponses rapides aux messages électroniques qu’il reçoit. Pour en savoir plus sur la manière dont ce système est censé fonctionner, consultez la page https://research.googleblog.com/2015/11/computer-respond-to-this-email.html. Dans le monde réel, il savait répondre dans la plupart des conversations par « I love you » car on lui avait appris à utiliser des exemples biaisés. Il s’est produit un phénomène similaire avec Tay, l’agent conversationnel de Twitter, de Microsoft, dont la capacité d’apprendre à partir des interactions avec les utilisateurs a fait qu’il s’est laissé dévoyer, certains utilisateurs lui ayant adressé des propos tendancieux ou malveillants.



Si vous voulez être au courant des dernières avancées en matière d’agents conversationnels, intéressez-vous aux concours annuels dans lesquels des tests de Turing sont appliqués à la technologie actuelle. Le prix Lobner, par exemple, en est le plus fameux (http://www.loebner.net/Prizef/loebner-prize.html) et c’est un bon exemple pour commencer. Le plus récent lauréat du prix Lobner, au moment où ces lignes sont écrites, est Mitsuku, un logiciel qui ne réussit pas le test de Turing mais qui est capable de raisonner sur des objets spécifiques proposés dans le discours. Il peut aussi jouer à des jeux, et même, exécuter des tours de magie (http://www.mitsuku.com/).

Mettre une IA en concurrence avec une autre IA

Les réseaux de neurones récurrents permettent à un ordinateur de converser avec vous, et si vous ne vous doutez pas que le réseau de neurones réactive des séquences de mots qu'il a appris, vous avez l'idée qu'il se passe en coulisses quelque chose qui relève de l'intelligence. En réalité, il n'y a là ni pensée ni raisonnement, même si la technologie ne fait pas que réutiliser des expressions présélectionnées.

Les réseaux antagonistes génératifs sont un autre type de technologie d'apprentissage profond qui peut vous donner davantage encore l'illusion que l'IA est capable de faire preuve de créativité. Là encore, cette technologie consiste à reprendre des exemples déjà stockés en mémoire, la machine interprétant les règles contenues dans ces exemples et s'en servant comme un enfant utilise des briques en plastique (techniquement, les règles sont les distributions statistiques sur lesquelles reposent les exemples). Les réseaux antagonistes génératifs n'en sont pas moins une technologie formidable ouvrant la voie à un grand nombre d'applications futures.

Les réseaux antagonistes génératifs sont apparus en 2014, issus des travaux d'un petit nombre de chercheurs du Département d'informatique et de recherche opérationnelle de l'université de Montréal, le plus remarquable d'entre eux étant Iam Goodfellow (voir le livre blanc à l'adresse https://arxiv.org/pdf/1406.2661.pdf). La nouvelle approche proposée, celle de l'apprentissage profond, a tout de suite suscité de l'intérêt et il s'agit à présent d'une des technologies auxquelles les chercheurs s'intéressent le plus. Elle fait constamment l'objet de développements et d'améliorations. Selon Yann LeCun,

les réseaux antagonistes génératifs seraient « l'idée la plus intéressante de ces dix dernières années dans le domaine de l'apprentissage machine ». Dans une interview accordée à la *MIT Technology Review*, Iam Goodfellow explique ainsi cet enthousiasme : « On peut considérer que les modèles génératifs donnent à l'intelligence artificielle une forme d'imagination » (https://www.technologyreview.com/lists/innovators-under-35/2017/inventor/ian-goodfellow/).

Pour voir fonctionner une version de base de réseau antagoniste génératif (il en existe maintenant un certain nombre de variantes perfectionnées, et davantage encore sont en cours de développement), il vous faut un jeu de données de référence, généralement constitué de données réelles, avec des exemples qu'il serait pertinent d'utiliser pour l'apprentissage. Si vous disposez d'images représentant des chiens, par exemple, le réseau devra apprendre, à partir de ces images, à quoi un chien ressemble. Ensuite, il pourra proposer des images plausibles et réalistes de chiens différents de ceux du jeu de données initial (de nouvelles images, sachant que la simple reproduction d'images existantes serait considérée comme une erreur).

Le jeu de données est le point de départ. Il vous faut aussi deux réseaux de neurones, chacun spécialisé dans une tâche différente, et ces deux réseaux doivent être mis en concurrence l'un avec l'autre. L'un, appelé le *générateur*, traite une entrée choisie arbitrairement (par exemple une séquence de nombres au hasard) et produit un résultat (par exemple l'image d'un chien), qui sera un artéfact puisqu'il aura été créé artificiellement. Le second réseau est le *discriminateur*, qui doit distinguer correctement les produits du générateur, c'est-à-dire les artéfacts, des exemples du jeu de données utilisé pour l'apprentissage.

Quand un réseau antagoniste génératif commence l'apprentissage, les deux réseaux essaient de progresser par rétropropagation, à partir des résultats du discriminateur. En essayant de distinguer une image réelle d'un artéfact, le discriminateur peut commettre une erreur (comme un réseau de neurones peut commettre une erreur au cours d'une tâche de classification). Les réponses correctes du discriminateur sont reçues par le générateur comme des erreurs (il n'a pas pu produire des artéfacts similaires aux images du jeu de données, et le discriminateur les a détectés). La Figure 11.5 illustre cette relation.

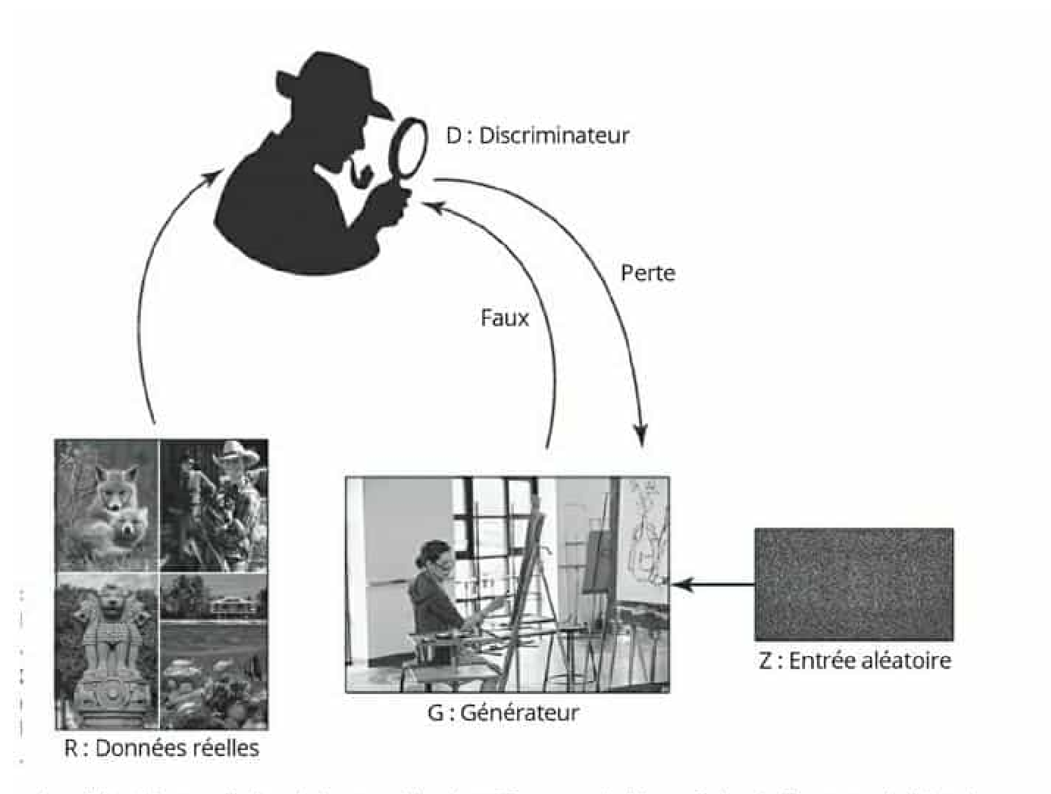


FIGURE 11.5 : Le fonctionnement d'un réseau antagoniste génératif, une oscillation entre un générateur et un discriminateur.

Les images originales choisies par Goodfellow pour expliquer comment fonctionne un réseau antagoniste

génératif sont celles du réalisateur de faux artistique et du détective. Le détective apprend à détecter les faux artistiques, mais le faussaire progresse également afin d'éviter que l'enquêteur détecte le faux.

Vous vous demandez peut-être comment le générateur apprend à créer les bons artéfacts s'il ne voit jamais un original. Seul le discriminateur voit le jeu de données original quand il s'efforce de distinguer l'œuvre originale des artéfacts du générateur. Même si le générateur n'examine jamais rien du jeu de données original, il reçoit des indices par le biais du travail du discriminateur. Ce sont de légers indices, le discriminateur étant guidé par les nombreux échecs initiaux du générateur. Tout se passe comme si vous deviez apprendre à peindre la Joconde sans l'avoir jamais vue, uniquement avec l'aide d'un ami qui vous guiderait dans vos suppositions. Cette situation rappelle le théorème sur l'armée infinie de singes qui tapent à la machine, mais il y a quelques différences. Selon ce théorème, on pourrait s'attendre à ce que les singes écrivent les poèmes de Shakespeare par pure chance (voir https://www.npr.org/sections/13.7/2013/12/10/249726951/the-infinite-monkey-theorem-comes-to-life). En l'occurrence, le générateur ne recourt à l'aléa qu'au départ, il est ensuite progressivement guidé par la rétroaction du discriminateur. Moyennant quelques modifications apportées à ce concept de base, les réseaux antagonistes génératifs sont devenus capables des prouesses suivantes :

> » Créer des images réalistes d'objets tels qu'articles de mode, décors d'intérieur ou design industriel à partir d'une description verbale (vous pouvez demander une fleur jaune et blanche, et vous l'obtenez, comme dans l'article suivant : https://arxiv.org/pdf/1605.05396.pdf)

» Modifier des images existantes en appliquant une résolution plus élevée, en ajoutant des motifs particuliers (par exemple, pour transformer un cheval en zèbre : https://junyanz.github.io/CycleGAN/) ou en remplissant des parties manquantes (si, par exemple, vous voulez faire disparaître une personne sur une photo, le réseau antagoniste génératif remplacera l'espace laissé vide par un morceau de décor plausible, comme dans cet exemple d'architecture neuronale : http://hi.cs.waseda.ac.jp/~iizuka/projects/ completion/en/).

» De nombreuses applications pionnières, comme la production de mouvements à partir de photos statiques, la création d'objets complexes comme des textes (on parle de prédiction structurée car le résultat n'est pas une simple réponse mais une série de réponses qui ont un lien entre elles), la création de données pour l'apprentissage machine supervisé, ou même la production d'une cryptographie élaborée (https://arstechnica.com/information-technology/2016/10/google-ai-neural-networkcryptography/)



Les réseaux antagonistes génératifs sont une technologie pionnière dans le domaine de l'apprentissage profond, et il existe divers domaines de recherche concernant son application en IA. Si l'IA doit un jour être dotée d'un pouvoir d'imagination et de création, ce sera probablement grâce à des technologies comme les réseaux antagonistes génératifs. Pour avoir une idée des progrès de ces technologies, lisez les pages consacrées à ces réseaux sur le site Internet d'OpenAI, un organisme de recherche sur l'IA à but non lucratif fondé par Greg Brockman, Ilya Sutskever, Elon Musk (qui a travaillé chez PayPal et

SpaceX et qui est le fondateur de Tesla) et Sam Altman (https://blog.openai.com/generative-models/).

PARTIE 4
Travailler avec l'IA dans des applications matérielles

DANS CETTE PARTIE…

» Travailler avec des robots

» Voler partout avec des drones

» Se laisser conduire par une IA

Chapitre 12

Mettre au point des robots

DANS CE CHAPITRE

» **Savoir distinguer les robots du monde réel de ceux de la science-fiction**

» **Réfléchir à une éthique de la robotique**

» **Trouver d'autres applications pour les robots**

» **Voir de quoi est fait un robot**

On confond trop souvent la robotique avec l'IA, bien qu'il s'agisse de deux choses nettement différentes. L'intelligence artificielle a pour objet de trouver des solutions à des problèmes difficiles liés aux capacités humaines (comme reconnaître les objets, ou comprendre un langage parlé ou écrit), tandis que la robotique consiste à utiliser des machines pour exécuter des tâches dans le monde physique de façon partiellement ou totalement automatisée. On peut considérer que l'IA est le logiciel utilisé pour trouver des solutions à des problèmes, tandis que la robotique est le matériel qui permet de faire de ces solutions une réalité.

Le matériel robotique peut fonctionner dans certains cas en utilisant du logiciel d'IA, et dans certains cas sans en utiliser. Certains robots sont commandés à distance par des humains, comme le robot da Vinci dont il est question dans la section « Assister un chirurgien » du Chapitre 7. D'autres robots exécutent de façon autonome des tâches

qui leur sont assignées sans qu'aucune intervention humaine soit nécessaire. Entre ces extrêmes, il y a des robots qui peuvent recevoir des humains des ordres abstraits (comme aller d'un point A à un point B sur une carte, ou aller saisir un objet) et exécuter ces ordres grâce à une IA. Dans bien des cas, l'IA constitue effectivement une augmentation, mais l'être humain garde le contrôle de l'activité. Quand une IA est intégrée à un robot, celui-ci est plus intelligent et plus utile pour l'exécution des tâches, mais les robots n'ont pas toujours besoin d'une IA pour fonctionner correctement. C'est l'imagination humaine qui a fait coïncider ces deux concepts, par le biais des films et des romans de science-fiction.

Ce chapitre analyse la manière dont s'est faite l'association entre les deux concepts, et établit une distinction entre les réalités actuelles en matière de robotique et les changements que les solutions à base d'IA sont susceptibles d'apporter. La fabrication de robots a commencé dans les années 1960. Ce chapitre étudie aussi l'utilisation accrue des robots dans les tâches industrielles, la recherche scientifique, la médecine et les guerres. Des découvertes récentes dans le domaine de l'IA accélèrent ce processus, sachant qu'elles permettent de résoudre des problèmes difficiles dans le domaine de la robotique, comme reconnaître des objets du monde réel, prédire le comportement humain, comprendre des commandes vocales, parler correctement, apprendre à marcher en position verticale, et plonger en arrière, mais oui, comme vous pouvez le lire dans cet article sur les progrès récents de la robotique : https://www.theverge.com/circuitbreaker/2017/11/17/16671328/boston-dynamics-backfliprobot-atlas.

Définir les rôles des robots

Les robots sont un concept relativement récent. Le mot vient du terme tchèque *robota*, qui signifie travaux forcés. Il est apparu pour la première fois en 1920 dans la pièce de théâtre *R.U.R.* (*Rossumovi univerzální roboti*) de Karel Čapek. Cependant, cela faisait déjà longtemps que les hommes rêvaient d'êtres mécaniques. On retrouve chez les anciens Grecs le mythe d'un homme mécanique en bronze, Talus, fabriqué par le dieu de la métallurgie Héphaïstos à la demande de Zeus, le père des dieux. Les mythes grecs comportent aussi des références à la fabrication par Héphaïstos d'autres *automates*. Les automates sont des machines qui fonctionnent toutes seules et exécutent des séquences de tâches spécifiques et prédéterminées (contrairement aux robots, qui sont doués d'une flexibilité leur permettant d'exécuter un vaste ensemble de tâches). Les Grecs avaient même fabriqué des automates hydrauliques dont le fonctionnement était analogue à celui d'un algorithme exécuté dans le monde réel. En tant qu'algorithmes, les automates intègrent l'intelligence de leur créateur, si bien qu'ils donnent l'illusion d'être des machines douées de raison et conscientes d'elles-mêmes.

On trouve des exemples d'automates en Europe dans la civilisation grecque, au Moyen Âge, à la Renaissance et dans les temps modernes. Au Moyen-Orient, le mathématicien et inventeur Al-Jazari a laissé un certain nombre de dessins (voir http://www.muslimheritage.com/article/al-jazari-mechanical-genius). La Chine et le Japon ont eu leurs propres versions des automates. Certains automates sont des élaborations mécaniques complexes, tandis que d'autres sont de vrais canulars, comme le Turc mécanique, une machine du XVIIIe siècle censée être capable de jouer aux échecs, mais dans laquelle un homme était caché.



Il importe de faire la différence entre les automates et les autres types d'animations à forme humaine. Le Golem, par exemple (https://www.myjewishlearning.com/article/golem/), était un mélange d'argile et de magie. Il ne comportait aucune machinerie, et ne relevait donc pas du type d'appareil étudié dans ce chapitre.

Les robots décrits par Čapek n'étaient pas exactement des automates mécaniques, c'étaient plutôt des êtres vivants conçus et fabriqués comme s'il s'agissait d'automates. Ils avaient une forme humaine et jouaient un rôle particulier dans la société, où ils étaient censés remplacer les ouvriers. Rappelant le Frankenstein de Mary Shelley, les robots de Čapek étaient un peu ce que l'on appellerait aujourd'hui des *androïdes* : des êtres artificiels issus des biotechnologies, comme ceux du roman de Philip K. Dick *Les androïdes rêvent-ils de moutons électriques ?* (dont s'inspire le film *Blade Runner*). Et cependant, le mot robot est aussi utilisé pour désigner des dispositifs mécaniques autonomes qui sont conçus non pas pour épater le public ni pour le distraire, mais pour produire des biens et des services. Par ailleurs, les robots sont devenus un thème central en science-fiction, que ce soit dans les livres ou les films, ce qui a incité davantage encore les gens à imaginer les robots comme des IA à forme humaine, conçues pour servir les humains, ce qui est assez proche de l'idée initiale des robots de Čapek. Peu à peu, cette idée est passée du domaine de l'art à celui des sciences et des technologies, et elle est devenue une source d'inspiration pour les chercheurs et les ingénieurs.



Čapek est à l'origine à la fois de l'idée des robots et de celle de l'apocalypse des robots, c'est-à-dire de la prise de contrôle de la société par l'IA comme dans certains films de science-fiction. Avec les progrès récents de l'IA, cette idée a fini par effrayer des personnalités éminentes

comme le fondateur de Microsoft, Bill Gates, le physicien Stephen Hawking et l'inventeur et homme d'affaires Elon Musk. À la fin de la pièce de Čapek, les esclaves robotiques se rebellent contre les humains qui les ont créés et exterminent presque toute l'humanité.

En finir avec la vision des robots présentée par la science-fiction

Le premier robot commercialisé, l'Unimate (https://www.robotics.org/joseph-engelberger/unimate.cfm) est apparu en 1961. C'était simplement un bras robotique : un bras mécanique programmable dont les liens et les articulations étaient en métal, avec une extrémité qui pouvait saisir, faire pivoter ou souder les objets manipulés selon des instructions définies par les opérateurs humains. General Motors l'avait acheté pour l'utiliser dans la production d'automobiles.

L'Unimate devait saisir des moulages sur la chaîne de montage et les souder, une tâche qui était physiquement dangereuse pour les ouvriers humains. Pour avoir une idée des capacités de cette machine, regardez cette vidéo : https://www.youtube.com/watch?v=hxsWeVtb-JQ. Les sections qui suivent décrivent les réalités actuelles dans le domaine des robots.

Étudier les lois de la robotique

Avant l'apparition d'Unimate et bien avant celle des autres bras robotiques utilisés sur les chaînes de montage, on savait déjà à quoi les robots ressembleraient, comment ils se comporteraient, et même, comment ils penseraient.

Isaac Asimov, écrivain américain connu pour ses ouvrages de science-fiction et de vulgarisation scientifique, avait publié dans les années 1950 une série de romans dans lesquels apparaissait un concept de robots complètement différent de ceux utilisés maintenant dans l'industrie.



C'est Asimov qui a inventé le mot *robotique*, et il l'a utilisé de la même manière qu'on utilise le mot *mécanique*. Son imagination fertile reste une référence aujourd'hui en matière de prédiction de ce que seront les robots. Asimov a mis en scène des robots à l'ère de l'exploration de l'espace, et ses robots utilisent leur cerveau positronique pour aider les humains au quotidien dans leurs tâches ordinaires et extraordinaires. Un *cerveau positronique* est un appareil fictif permettant aux robots, dans les romans d'Asimov, d'agir de façon autonome et d'assister ou de remplacer les humains dans un certain nombre de tâches. Outre ses capacités comparables à celles des humains en matière de compréhension et d'action (IA forte), le cerveau positronique se conforme aux Trois lois de la robotique, qui sont incorporées à la machine et font que le comportement des robots est moral :

1. **Un robot ne peut porter atteinte à un être humain, ni, en restant passif, permettre qu'un être humain soit exposé au danger.**
2. **Un robot doit obéir aux ordres qui lui sont donnés par un être humain, sauf si de tels ordres entrent en conflit avec la première loi.**
3. **Un robot doit protéger son existence tant que cette protection n'entre pas en conflit avec la première ou la deuxième loi.**

   Par la suite, l'auteur a ajouté une loi zéro ayant priorité sur les autres lois afin d'assurer que les robots agissent en faveur de la sécurité du plus grand nombre :

4. **Un robot ne peut porter atteinte à l'humanité, ni, en restant passif, permettre que l'humanité soit exposée au danger.**

Grâce à ces Trois lois, qui jouent un rôle central dans toutes les histoires de robots d'Asimov, les robots peuvent travailler avec les humains sans aucun risque de rébellion ou d'apocalypse de l'IA. Impossibles à contourner et à modifier, ces trois lois s'exécutent par ordre de priorité et se présentent comme des formulations mathématiques dans les fonctions du cerveau positronique. Malheureusement, ces lois ne sont pas exemptes de failles ni d'ambiguïté, et c'est là ce qui nourrit l'intrigue de la plupart des romans d'Asimov. Les Trois lois sont censées provenir d'un *Manuel de robotique*, 56e édition, publié en 2058 après J.-C., et elles reposent sur des principes d'innocence, d'obéissance et de survie.

Asimov a imaginé un univers dans lequel la morale peut se réduire à un petit nombre de principes simples, avec quelques risques qui donnent la matière de l'intrigue dans de nombreuses histoires. En réalité, Asimov voyait les robots comme des outils et pour lui, les Trois lois étaient viables même dans le monde réel pour contrôler leur utilisation (pour plus de détails, lisez cette interview accordée en 1981 à *Compute ! Magazine* : https://archive.org/stream/1981-11-compute-magazine/Compute_Issue_018_1981_Nov#page/n19/mode/2up). Cependant, à l'encontre de la vision optimiste d'Asimov, les robots actuels ne possèdent pas les capacités suivantes :

- comprendre les trois lois de la robotique ;
- choisir des actions en fonction de ces trois lois ;
- détecter et reconnaître une violation possible de ces trois lois.

On pourrait penser que les robots d'aujourd'hui ne sont pas très intelligents, en réalité, puisqu'il leur manque ces capacités. Or, l'Engineering and Physical Sciences Research Council (EPSRC), qui est au Royaume-Uni le principal organisme de financement de la recherche en ingénierie et sciences physiques, a proposé en 2010 de réviser les lois de la robotique d'Asimov dans le cadre de l'utilisation des robots réels, compte tenu de la technologie du moment. Le résultat est très différent des énoncés originaux d'Asimov (voir https://www.epsrc.ac.uk/research/ourportfolio/themes/engineering/activities/principlesofrobotics/). Selon ces principes révisés, il est admis que les robots peuvent même tuer (pour des raisons de sécurité nationale) puisqu'ils sont des outils. Comme dans l'utilisation des autres outils, c'est à l'utilisateur humain de respecter la loi et de se conformer aux principes de la morale, et non à la machine qui ne fait qu'exécuter. En outre, il faut toujours que quelqu'un (un être humain) ait des comptes à rendre concernant les résultats des actions d'un robot.



Les principes de l'EPSRC représentent un point de vue plus réaliste sur les robots et la morale, sachant que les technologies actuelles reposent sur l'IA faible, mais ils pourraient aussi constituer une solution partielle dans des scénarios de technologies plus avancées. Le Chapitre 14 étudie des problèmes relatifs à l'utilisation de voitures sans conducteur, avec une sorte de robot mobile qui conduit à votre place. Dans l'étude du dilemme du tramway, par exemple, plus loin dans ce chapitre, on se retrouve confronté à des problèmes de morale qui remettent en question la pertinence qu'il y aurait à compter sur des machines automatiques lorsque certains choix doivent être faits.

Définir les capacités réelles des robots

Non seulement les capacités existantes des robots sont encore loin d'égaler celles des robots humanoïdes que l'on trouve dans les œuvres d'Asimov, mais ces robots appartiennent aussi à des catégories différentes. Le genre de robots bipèdes imaginés par Asimov est actuellement ce qu'il y a de plus rare et de moins évolué.

La catégorie de robots la plus fréquente est le bras robotique, comme Unimate qui a été présenté précédemment. Les robots de cette catégorie sont aussi appelés des *robots manipulateurs* ou des *bras manipulateurs*. On en voit dans les usines, où ils sont utilisés dans des activités industrielles : ils assemblent et soudent à une vitesse et avec une précision dont les ouvriers humains ne seraient pas capables. Certains bras manipulateurs sont aussi utilisés dans les hôpitaux pour aider aux opérations chirurgicales. Les bras manipulateurs ont une amplitude de mouvement limitée car ils sont fixés quelque part (ils peuvent parfois se déplacer, mais très peu car soit ils ne sont pas équipés d'un moteur puissant, soit ils ont besoin d'un branchement électrique), et leur déménagement ne peut se faire qu'avec l'aide de techniciens spécialisés. Par ailleurs, en général les robots manipulateurs utilisés pour la production industrielle sont entièrement automatisés (contrairement aux appareils utilisés en chirurgie qui sont télécommandés par le praticien, les décisions étant prises par celui-ci). Plus d'un million de robots manipulateurs sont utilisés dans le monde, la moitié au Japon.

La deuxième catégorie de robots la plus importante, et qui connaît actuellement un essor, est celle des robots mobiles. Leur spécialité, contrairement aux robots manipulateurs, consiste à se déplacer au moyen de roues, de rotors, d'ailes, ou même de jambes. Dans cette vaste

catégorie, on trouve des robots qui livrent de la nourriture (https://nypost.com/2017/03/29/dominos-delivery-robots-bring-pizza-to-the-final-frontier/) ou des livres (https://www.digitaltrends.com/cooltech/amazon-prime-air-delivery-drones-history-progress/) aux entreprises, et même des robots qui explorent la planète Mars (https://mars.nasa.gov/mer/overview/). Les robots mobiles sont généralement sans conducteur (personne ne voyage avec eux) et commandés à distance, mais ils deviennent de plus en plus autonomes et l'on peut s'attendre à voir apparaître dans cette catégorie des robots plus indépendants. Il existe deux types particuliers de robots mobiles : les robots volants, ou drones (Chapitre 13), et les voitures sans conducteur (Chapitre 14).

La dernière catégorie de robots est celle des *robots manipulateurs mobiles*, qui peuvent se déplacer (comme les autres robots mobiles) et manipuler (comme les bras robotiques). Dans cette catégorie, le sommet, ce n'est pas simplement un robot qui se déplace et qui est pourvu d'un bras mécanique, mais un robot qui imite aussi la forme et le comportement des êtres humains. Le *robot humanoïde* est un bipède (il se déplace sur deux jambes), il a un torse humain et il communique avec les humains par la voix et par des expressions. Ce sont les robots que la science-fiction a imaginés, mais ils ne sont pas faciles à mettre au point.

Savoir pourquoi il est difficile d'être un humanoïde

Les robots humanoïdes sont difficiles à mettre au point, et les chercheurs y travaillent toujours. Un robot humanoïde

nécessite des capacités renforcées en IA pour pouvoir être autonome, mais il faut aussi qu'il puisse se déplacer de la même manière que nous. Cependant, le plus gros problème, c'est de faire accepter aux humains une machine qui leur ressemble. Les sections qui suivent abordent plusieurs aspects de la création d'un robot humanoïde.

Créer un robot qui marche

Réfléchissons au problème que pose la conception d'un robot marchant sur deux jambes (un *robot bipède*). C'est une chose que nous apprenons à pratiquer avec agilité et sans y penser, mais qui est très problématique pour un robot. Un robot à quatre pieds n'a pas de difficulté à se tenir en équilibre et ne consomme pas beaucoup d'énergie pour cela. L'être humain, en revanche, consomme de l'énergie rien que pour rester debout, ainsi que pour maintenir son équilibre et pour marcher. Un robot humanoïde, comme nous, doit se tenir continuellement en équilibre, et il faut qu'il y parvienne d'une manière qui soit à la fois efficace et économique. Or, il faut pour cela qu'il soit équipé d'une batterie volumineuse, lourde et encombrante, ce qui complique encore le problème de l'équilibre.

Une vidéo proposée par IEEE Spectrum donne une idée plus précise de la difficulté que peut représenter le simple fait de marcher. Elle montre des robots qui participent au défi robotique de la DARPA, un projet organisé entre 2012 et 2015 sous les auspices de l'Agence des projets de recherche avancée du ministère américain de la Défense : https://www.youtube.com/watch?v=g0TaYhjpOfo. L'objet de ce défi était d'étudier les avancées dans le domaine de la robotique susceptibles de permettre un progrès dans les opérations de sauvetage en cas de catastrophe dans des environnements qui sont

dangereux pour les humains (https://www.darpa.mil/program/darpa-robotics-challenge). Il s'agit donc de robots qui marchent sur différents terrains, qui ouvrent des portes, qui saisissent des outils, par exemple une perceuse électrique, ou qui essaient de manipuler un volant de commande de valve. Un robot conçu récemment par Boston Dynamics et nommé Atlas semble prometteur, comme l'explique cet article : https://www.theverge.com/circuitbreaker/2017/11/17/16671328/boston-dynamics-backfliprobot-atlas. Le robot Atlas est vraiment exceptionnel, mais il reste encore un long chemin à parcourir.



Un robot avec des roues peut facilement se déplacer sur des routes, mais dans certaines situations, pour répondre à des besoins particuliers, c'est d'un robot humanoïde dont on a besoin. La plupart des infrastructures, dans le monde, sont conçues pour être franchies par des hommes et des femmes. La présence d'obstacles, la largeur d'un couloir ou la présence de portes ou de marches sont un problème pour des robots ayant une morphologie différente. En cas d'urgence dans une centrale nucléaire, par exemple, il faudrait un robot capable de pénétrer dans l'enceinte du réacteur pour fermer une valve. C'est en ayant une forme humaine qu'il pourra circuler dans les installations, descendre des marches et aller tourner le volant actionnant la valve.

Surmonter les réticences des humains : la vallée dérangeante

Les robots humanoïdes qui ressemblent un peu trop aux humains posent un problème. En 1970, Masahiro Mori, professeur à l'Institut de technologie de Tokyo, a étudié

l'impact des robots sur la société japonaise. Il a inventé le terme *Bukimi no Tani Genshō*, que l'on peut traduire par *vallée dérangeante*. Mori a constaté que lorsque les robots deviennent plus réalistes, les humains se sentent davantage en affinité avec eux, mais que cette progression du degré de familiarité n'est vérifiée que si les robots n'ont pas atteint un certain degré de réalisme, au-delà duquel les gens se mettent à ne plus les aimer du tout (et même, à en éprouver de la répulsion). Cette répulsion s'accentue jusqu'à ce que le robot atteigne le niveau de réalisme qui le rend tout à fait semblable à un être humain. Cette progression est représentée sur la Figure 12.1 et décrite dans l'article original de Mori, à l'adresse https://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley.

Différentes hypothèses ont été formulées concernant les raisons de cette répulsion qu'éprouvent les gens vis-à-vis d'un robot qui est presque humain sans l'être entièrement. Les indices utilisés par les humains pour détecter les robots sont le ton de la voix, la rigidité des mouvements et la texture artificielle de la peau. Certains scientifiques attribuent la vallée dérangeante à des facteurs culturels, d'autres à des facteurs psychologiques ou biologiques. Une expérience récente sur des singes montre que ces derniers vivent une expérience similaire face à des photos plus ou moins réalistes de singes modifiées par une technologie 3D (voir l'article : https://www.wired.com/2009/10/uncanny-monkey/). Les singes qui ont participé à cette expérience ont manifesté une légère aversion aux photos réalistes, ce qui suggère une raison biologique commune à la vallée dérangeante. L'explication pourrait donc tenir à une réaction d'autoprotection face à des êtres perçus négativement comme non naturels parce qu'ils sont malades, ou peut-être même morts.

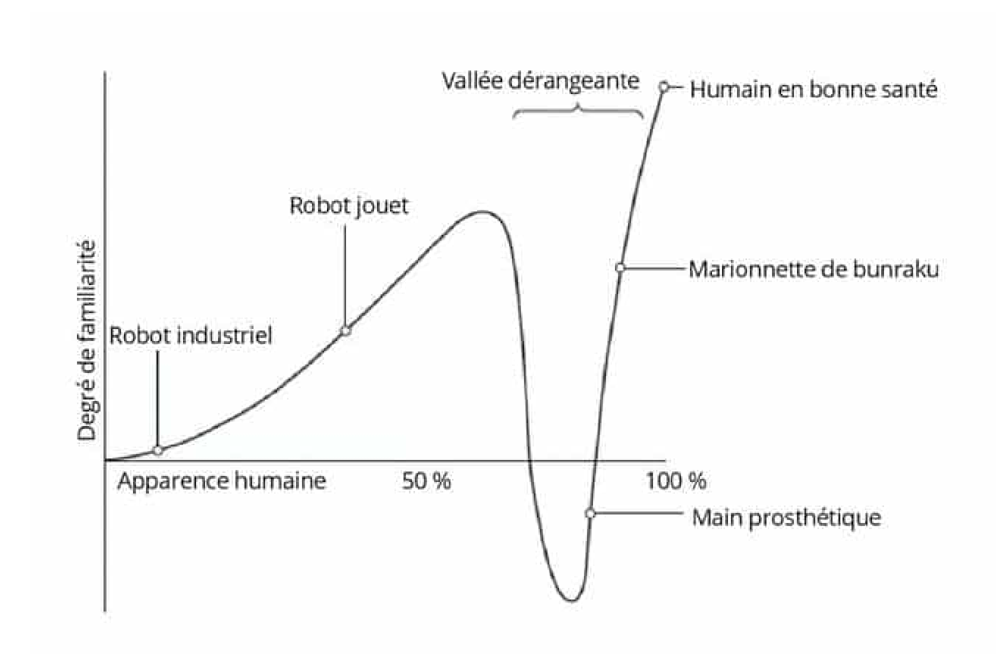


FIGURE 12.1 : La vallée dérangeante.

Ce qui est intéressant ici, c'est de constater que si nous avons besoin de robots humanoïdes pour fournir une assistance à des personnes, nous devons aussi étudier leur degré de réalisme et certains détails esthétiques essentiels, afin de pouvoir obtenir une réponse émotionnelle positive grâce à laquelle les utilisateurs accepteront l'aide du robot. De récentes observations montrent que même des robots présentant peu de ressemblance avec l'être humain suscitent des sentiments affectifs et créent des liens avec leurs utilisateurs. Ainsi, de nombreux soldats américains évoquent un sentiment de perte quand leurs petits robots tactiques conçus pour la détection et le maniement d'explosifs sont détruits au cours de leur mission (vous pouvez lire un article à ce sujet dans la revue technique du MIT : https://www.technologyreview.com/s/609074/how-we-feel-about-robots-that-feel/).

Travailler avec des robots

Différents types de robots ont des applications différentes. À mesure qu'ont été élaborés et améliorés les robots des trois catégories (manipulateurs, mobiles, et humanoïdes), de nouveaux champs d'application se sont offerts à la robotique. Il est aujourd'hui impossible d'énumérer de manière exhaustive toutes les utilisations existantes de robots, mais les sections qui suivent abordent les utilisations les plus prometteuses et les plus révolutionnaires.

Améliorer la performance économique

Les robots manipulateurs, ou robots industriels, représentent toujours le plus gros pourcentage de robots en activité dans le monde. D'après World Robotics 2017, une étude préparée par la Fédération internationale de la robotique, plus de 1800000 robots étaient en activité dans l'industrie à la fin de l'année 2016 (lire ici un résumé de cette étude : https://ifr.org/downloads/press/Executive_Summary_WR_2017_Industrial_Robots.pdf). Le nombre de robots industriels devrait atteindre 3 millions en 2020 en raison du boom de l'automation dans ce secteur. En fait, les usines (en tant qu'entités) utiliseront des robots pour devenir plus intelligentes, un concept désigné par l'appellation *Industry 4.0*. Grâce à l'utilisation généralisée de l'Internet, des capteurs, des données et des robots, les solutions Industry 4.0 permettent d'obtenir une personnalisation plus facile et une meilleure qualité des produits en moins de temps que ce qui serait possible sans les robots. Toujours est-il que des robots sont déjà en activité dans des environnements dangereux, et pour des tâches comme la soudure, l'assemblage, la peinture et l'emballage, qu'ils travaillent plus vite, avec une plus

grande précision, et pour des coûts moindres que les ouvriers humains.

Prendre soin de la personne

Depuis 1983, des robots assistent les chirurgiens dans des opérations difficiles en pratiquant des incisions précises et exactes comme seuls des bras robotiques en sont capables. En dehors de la possibilité d'un contrôle à distance des opérations (le chirurgien restant à l'extérieur du bloc opératoire afin que l'environnement du patient soit plus stérile), les progrès dans le domaine de l'automatisation des opérations ouvrent la voie à la possibilité d'actes chirurgicaux entièrement automatisés dans un proche avenir, comme l'annonce cet article : https://www.huffingtonpost.com/entry/is-thefuture-of-robotic-surgery-already-here_us_58e8d00fe4b0acd784ca589a.

Dispenser des services

Les robots fournissent d'autres types de services, dans le secteur privé comme dans le secteur public. Le plus fameux robot d'intérieur est l'aspirateur Roomba, un robot qui passe l'aspirateur tout seul dans votre maison (c'est un grand succès commercial dans le domaine de la robotique, avec plus de 3 millions d'exemplaires vendus), mais il convient de ne pas oublier d'autres exemples de services assurés par des robots :

- » **Les livraisons :** un exemple est le robot de Domino's pizza (*https://www.youtube.com/watch?v=PL-4K4yqWHE*).
- » **La tonte de la pelouse :** il existe une variété incroyable de tondeuses robotisées pour le jardin : vous en trouverez dans les magasins de jardinage.

» **Les informations et les divertissements :** un exemple est Pepper, que l'on trouve dans tous les magasins SoftBank au Japon (http://mashable.com/2016/01/27/softbank-pepper-robot-store/).

» **Les soins aux personnes âgées :** un exemple de robot dispensant des soins aux personnes âgées est le robot Hector, un projet financé par l'Union européenne (https://www.forbes.com/sites/jenniferhicks/2012/08/13/hector-robotic-assistance-for-the-elderly/#5063a3212443).

Les robots d'assistance aux personnes âgées sont loin d'offrir une assistance générale comparable à celle des infirmières. Les robots assurent des tâches critiques comme rappeler quels sont les médicaments à prendre, aider les patients à passer du lit au fauteuil roulant, vérifier l'état physique du patient, déclencher une alarme en cas de problème, ou simplement jouer le rôle d'un compagnon. Le robot thérapeutique Paro, par exemple, dispense une thérapie animale aux personnes âgées à mobilité réduite, comme l'explique cet article : https://www.huffingtonpost.com/the-conversation-global/robot-revolution-why-tech_b_14559396.html.

S'aventurer dans des environnements dangereux

Les robots vont là où les gens ne peuvent pas aller, ou là où il serait très dangereux de se rendre. Certains robots ont été envoyés dans l'espace (les robots rover de la NASA Opportunity et Curiosity, envoyés sur Mars, en sont

les exemples les plus remarquables), et d'autres robots participeront aux futures explorations spatiales (le Chapitre 16 traite des robots envoyés dans l'espace). Beaucoup d'autres robots restent sur terre et sont utilisés pour des tâches souterraines comme transporter le minerai dans les mines ou produire des cartes des tunnels. Certains robots explorent même les égouts, notamment Luigi (un nom inspiré du frère d'un célèbre plombier dans des jeux vidéo). Luigi est un robot dragueur d'égouts conçu par le laboratoire Senseable City Lab du MIT pour étudier la santé publique là où les humains ne peuvent pas rester sains et saufs en raison de fortes concentrations de produits chimiques, de bactéries et de virus (voir http://money.cnn.com/2016/09/30/technology/mit-robots-sewers/index.html).

Des robots sont même utilisés là où les humains mourraient fatalement, par exemple dans le contexte de catastrophes nucléaires comme celles de Three Mile Island, de Tchernobyl et de Fukushima. Ces robots retirent les matières radioactives et rendent le périmètre moins dangereux. Les radiations à haute dose affectent même les robots, sachant qu'elles provoquent du bruit électronique et des fluctuations du signal qui, à la longue, endommagent les circuits. Les robots doivent être équipés de *composants électroniques résistants aux rayonnements* de manière à résister suffisamment aux effets des radiations pour pouvoir accomplir leur mission. C'est le cas de Little Sunfish, un robot submersible qui intervient dans un des réacteurs inondés de Fukushima, là où la fusion s'est produite (lire l'article à l'adresse http://www.bbc.com/news/in-pictures-40298569).

Les guerres et les scènes de crimes sont aussi des situations dangereuses, et des robots sont souvent utilisés pour transporter des armes ou désamorcer des bombes.

Ces robots peuvent aussi vérifier des paquets susceptibles de contenir des objets dangereux autres que les bombes. Des modèles de robots comme PackBot, d'iRobot (la société qui fabrique Roomba, l'aspirateur robot) ou Talon, de QinetiQ North America, manipulent des explosifs dangereux, commandés à distance par un spécialiste. Certains robots peuvent même remplacer des soldats ou des policiers dans des missions de reconnaissance ou pour des interventions directes (la police de Dallas, par exemple, a utilisé un robot pour abattre un tueur http://edition.cnn.com/2016/07/09/opinions/dallas-robot-questions-singer/index.html).



On s'attend à ce que les militaires utilisent de plus en plus les robots dans l'avenir. Au-delà des considérations éthiques concernant ces nouvelles armes, il s'agit d'une nouvelle version de l'ancien arbitrage « le beurre ou les canons » (https://www.huffingtonpost.com/jonathan-tasini/guns-versusbutter-our-re_b_60150.html), dans le sens où un pays peut renoncer à un certain pouvoir économique en échange d'une force militaire. Les robots semblent parfaitement adaptés à ce modèle, surtout concernant les armements traditionnels qui ne peuvent être maniés que par un personnel formé et entraîné. Ainsi, grâce à son rendement productif, un pays peut se doter à tout moment d'une armée de robots performante, une chose que le film *La Guerre des étoiles* annonce très clairement.

Comprendre le rôle des robots spécialisés

Les robots spécialisés, ce sont par exemple les drones et les voitures sans conducteur. Les drones posent un problème en raison de leur utilisation possible dans les

conflits armés, mais on utilise aussi des véhicules aériens sans pilote pour la surveillance, l'agriculture et bien d'autres activités moins inquiétantes, comme l'explique le Chapitre 13.

Les gens fantasment depuis longtemps sur les voitures sans conducteur. Ces voitures sont en train de devenir rapidement une réalité, après les exploits du Grand Challenge de la DARPA. Les constructeurs automobiles se sont généralement rendu compte que s'ils pouvaient produire et commercialiser des automobiles qui se conduisent toutes seules, cela pourrait entraîner un changement dans l'équilibre économique mondial (d'où cette course pour parvenir le plus vite possible à commercialiser un véhicule viable : https://www.washingtonpost.com/news/innovations/wp/2017/11/20/robot-drivenubers-without-a-human-driver-could-appear-as-early-as-2019/). Le Chapitre 14 étudie plus en détail les voitures sans conducteur, leur technologie et leurs implications.

Assembler un robot basique

Ce panorama des robots ne serait pas complet sans une étude de la manière d'en fabriquer un, en tenant compte de l'état actuel des avancées technologiques, et des améliorations qu'une IA peut apporter à son fonctionnement. Les sections qui suivent traitent des bases de la robotique.

Décliner les composants

Un robot étant conçu pour agir dans le monde, il lui faut des *effecteurs*, c'est-à-dire des jambes mobiles ou des roues lui permettant de se déplacer. Il lui faut aussi des

bras et des pinces pour attraper, faire pivoter et orienter des objets, et exercer ainsi des fonctions de *manipulation*. Concernant les capacités d'un robot, on parle aussi d'*actionneurs*, un mot utilisé comme synonyme d'effecteurs. L'actionneur est un des mécanismes dont sont constitués les effecteurs, il permet un mouvement unique. La jambe d'un robot comporte donc différents actionneurs, à savoir, des moteurs électriques ou des vérins hydrauliques qui exécutent des mouvements comme orienter le pied ou fléchir le genou.

Pour agir dans le monde, il faut déterminer la composition du monde et comprendre où réside le robot dans le monde. Les *capteurs* fournissent au robot des informations sur ce qui se produit autour de lui. Des dispositifs comme les caméras, les lidars, les sonars et les capteurs de pression mesurent l'environnement et indiquent au robot ce qui se passe et quelle est sa localisation. Le robot est donc avant tout un ensemble organisé de capteurs et d'effecteurs. Tous les éléments sont conçus pour travailler ensemble selon une certaine architecture, laquelle est précisément ce qui fait le robot (les capteurs et les effecteurs n'étant que des composants mécaniques et électroniques pouvant être utilisés séparément dans différentes applications.)

L'architecture interne commune est constituée de processus parallèles rassemblés en couches spécialisées chacune dans la résolution d'un type de problème. Le parallélisme est un aspect important. Nous autres humains, nous percevons un unique flux de conscience et d'attention : nous n'avons pas besoin de penser à ces fonctions essentielles que sont la respiration, les pulsations cardiaques et la digestion, car ces processus s'accomplissent d'eux-mêmes parallèlement à la pensée consciente. Souvent, nous pouvons même exécuter une action, comme marcher ou conduire, tout en parlant ou en faisant autre chose (quoique cela puisse se révéler

dangereux dans certaines situations). Il en est de même des robots. Dans une architecture en trois couches, par exemple, les divers processus du robot sont regroupés selon trois niveaux, chacun caractérisé par un temps de réponse différent et un degré de complexité de réponse différent :

- » **Le niveau réactif :** le système reçoit immédiatement les données transmises par les capteurs et par les canaux permettant au robot de percevoir le monde, et il réagit immédiatement aux problèmes soudains (par exemple, passé un angle, il tourne tout de suite pour ne pas entrer en collision avec un mur inconnu).
- » **Le niveau exécutif :** le système traite les données transmises par les capteurs, détermine où se trouve le robot dans le monde (il s'agit d'une fonction importante appelée la *localisation*), et décide quelle action exécuter compte tenu des exigences de la couche précédente, la couche réactive, et de la couche suivante, la couche délibérative.
- » **Le niveau délibératif :** le système planifie les tâches à exécuter, par exemple la manière dont le robot doit se rendre d'un point à un autre, et décide quelle séquence d'actions il doit effectuer pour aller saisir un objet. Cette couche définit pour le robot une série d'exigences que la couche exécutive va devoir satisfaire.

Une autre architecture souvent utilisée est l'architecture à pipeline, notamment pour les automobiles sans conducteur. Elle consiste à simplement diviser les processus parallèles du robot en phases distinctes comme la détection, la perception (qui consiste à comprendre ce que l'on a détecté), la planification, et le contrôle.

Détecter le monde

Le Chapitre 14 traite en détail des capteurs et présente des applications pratiques pour expliquer le fonctionnement des voitures sans conducteur. Il existe divers types de capteurs, certains orientés vers le monde extérieur, et d'autres sur le robot lui-même. Un bras robotique, par exemple, a besoin de savoir sur quelle distance il s'est déplié et s'il a atteint ou non la limite de son extension. Par ailleurs, certains capteurs sont actifs (ils recherchent activement l'information en fonction d'une décision du robot), tandis que d'autres sont passifs (ils ne font que recevoir l'information de façon permanente). Chaque capteur fournit une entrée électronique que le robot peut immédiatement exploiter ou traiter afin d'acquérir une perception.

La *perception* consiste à créer une carte locale des objets du monde réel et à déterminer la localisation du robot sur une carte plus générale du monde connu. En combinant les données provenant des différents capteurs, un processus appelé la *fusion des données* (ou des capteurs), le robot crée une liste de faits élémentaires qu'il va exploiter. L'apprentissage machine est utile ici, s'agissant d'obtenir des algorithmes de vision qui utilisent l'apprentissage profond pour reconnaître les objets et segmenter les images (voir Chapitre 11). Il s'agit également de rassembler toutes les données pour obtenir une représentation significative, à l'aide d'algorithmes d'apprentissage machine non supervisé. Cette tâche est ce que l'on appelle l'*intégration à faible dimension*, c'est-à-dire la traduction de données complexes provenant des capteurs sous forme d'une simple carte à deux dimensions ou d'une autre représentation. La détermination de la localisation du robot est ce que désigne l'expression *cartographie et localisation simultanées* (SLAM). Tout se passe comme lorsque vous

observez une carte pour savoir où vous vous trouvez dans une ville.

Contrôler un robot

Lorsque la détection a permis d'obtenir les informations nécessaires, la planification consiste à établir la liste des actions correctes à exécuter par le robot pour atteindre ses objectifs. La planification se fait par programmation (en utilisant par exemple un système expert, voir Chapitre 3) ou par un algorithme d'apprentissage machine comme les réseaux bayésiens (voir Chapitre 10). Les développeurs expérimentent actuellement le recours à l'apprentissage par renforcement (un apprentissage machine fondé sur le principe des essais et erreurs), mais un robot n'est pas un petit enfant (qui fonctionne aussi par essais et erreurs pour apprendre à marcher) : cette expérimentation risque de se révéler coûteuse en temps, frustrante et onéreuse lorsqu'il s'agit de créer automatiquement un plan, car le robot risque de subir des dommages au cours du processus.

Enfin, la planification n'est pas simplement l'affaire d'algorithmes intelligents, car dans la phase d'exécution les choses ne se déroulent généralement pas comme prévu. Envisageons le problème de notre point de vue d'êtres humains. Si l'on vous bande les yeux, même si vous désirez avancer droit devant vous, vous ne le ferez que si vous disposez d'une source constante de correctifs. Le résultat est que vous allez progresser en louvoyant. Vos jambes, qui sont les actionneurs, n'exécutent pas toujours parfaitement les instructions. Les robots sont confrontés au même problème, et aussi à d'autres problèmes comme des retards dans le processus (techniquement, on parle de *latence*) ou le fait que le robot n'exécute pas les instructions exactement à temps, ce qui entraîne un cafouillage. Cependant, le plus souvent, le

problème est lié à l'environnement du robot et relève d'un des deux cas de figure suivants :

- **Incertitude :** le robot n'est pas sûr de savoir où il est, ou bien il peut observer partiellement la situation mais sans la caractériser précisément. En raison de l'incertitude, les développeurs disent que le robot fonctionne dans un environnement stochastique.
- **Situations conflictuelles :** des personnes ou des objets en mouvement sont sur le passage. Dans certaines situations, ces objets peuvent même devenir. C'est un problème multiagent.



Les robots doivent fonctionner dans des environnements qui sont en partie inconnus, la plupart du temps imprédictibles et en perpétuel changement, ce qui signifie que toutes les actions devant s'enchaîner, ils doivent continuellement gérer un flux d'informations et d'actions en temps réel. On ne peut pas entièrement prédire ni programmer l'adaptation à ce genre d'environnement, et cette adaptation suppose des capacités d'apprentissage dont les algorithmes d'IA, de plus en plus, dotent les robots.

Chapitre 13
Voler avec des drones

DANS CE CHAPITRE

» **Faire la distinction entre les drones à usage militaire et civil**

» **Découvrir les utilisations possibles des drones**

» **Déterminer ce que l'IA peut permettre aux drones d'accomplir**

» **Prendre acte de la réglementation et des limites de l'exploitabilité des drones**

Les drones sont des robots mobiles qui se déplacent dans l'environnement en volant. Initialement liés à la guerre, les drones sont devenus une innovation remarquable dans les activités de loisirs, d'exploration, de livraison, *etc*. Cependant, les innovations dans le domaine militaire se poursuivent, et de nombreux spécialistes de l'IA aussi bien que des personnalités publiques se demandent si les robots ne risquent pas devenir un de ces jours des machines à tuer que rien ne pourrait arrêter.

Il est un rêve que les hommes ont réalisé le jour où les frères Wright ont volé pour la première fois le 17 décembre 1903 (voir https://www.nps.gov/wrbr/learn/historyculture/thefirstflight.htm). Les hommes ont toujours eu envie d'y parvenir, et des penseurs légendaires comme Léonard de Vinci, ce génie de la Renaissance, ont participé à des efforts dans ce sens (pour plus de détails, lisez cet article du Smithsonian Museum :

https://airandspace.si.edu/stories/editorial/leonardo-da-vinci-and-flight). Les technologies permettant de voler ont bien progressé. Les drones ont atteint une plus grande maturité que les autres robots mobiles, car la technologie qui leur permet de voler est bien maîtrisée. Ce qui change tout, c'est l'intégration de l'IA. Le déplacement dans les airs implique d'importantes limites à ce que les drones peuvent faire, notamment la charge qu'ils peuvent transporter et les actions qu'ils peuvent accomplir quand ils arrivent à destination.

Ce chapitre traite de la situation présente en matière de drones, qu'ils soient destinés au grand public, à un usage commercial ou à une application militaire. Il étudie aussi les rôles que les drones pourraient jouer dans l'avenir. Ces rôles dépendent en partie de l'intégration avec des solutions d'IA qui leur donneront davantage d'autonomie et des capacités accrues en termes de déplacement et de fonctionnement.

Prendre connaissance de l'état actuel des progrès

Les drones sont des robots mobiles qui volent et ils existent depuis longtemps, surtout dans les applications militaires (c'est de là que provient leur technologie). Ces machines volantes sont plus couramment appelées « drones » parce que leur bruit fait penser au bourdonnement des abeilles mâles, mais dans les documents officiels on les appelle plutôt systèmes d'aéronefs sans pilote ou UAS (*Unmanned Aircraft System*), ou avions de combat sans pilote, ou encore, aéronefs téléguidés.



Les noms sont parfois très parlants. Cet article d'*ABCNews* vous présente les acronymes et les noms officiels par lesquels on désigne les drones en anglais : http://www.abc.net.au/news/2013-03-01/drone-wars-the-definition-dogfight/4546598.

Envoyer des engins sans pilote dans des missions

Les drones militaires ont des ailes et une ou plusieurs hélices (ou des réacteurs) et ils ressemblent assez aux avions (bien que généralement plus petits). Ils présentent même souvent une certaine ressemblance avec les avions civils. Les drones militaires en sont déjà à leur sixième génération (voir https://www.military.com/daily-news/2015/06/17/navy-airforce-to-develop-sixth-generation-unmanned-fighter.html). Ils sont contrôlés à distance à l'aide de communications par satellite, parfois même de l'autre côté de la terre. L'opérateur accède aux informations par télémétrie et par la vision transmise par le drone qu'il contrôle, et ces informations lui permettent d'adresser à la machine des commandes particulières. Certains drones assurent des tâches de surveillance et de reconnaissance : ils sont simplement munis de caméras ou autres dispositifs d'acquisition d'informations. D'autres sont équipés d'armes et peuvent effectuer des attaques meurtrières sur des objectifs déterminés. Les modèles les plus redoutables ont des capacités comparables à celles des avions pilotés et peuvent atteindre n'importe quel point sur la terre, même des endroits où un pilote ne pourrait pas facilement se rendre.

Les drones militaires ont une longue histoire. Il est difficile de dire à quel moment cette histoire commence, mais la

Royal Navy utilisait déjà des aéronefs similaires aux drones pour le tir sur cibles dans les années 1930. Les États-Unis ont commencé à utiliser de vrais drones de façon régulière dès 1945 pour le tir sur cibles (voir http://www.designation-systems.net/dusrm/m-33.html). En 1971, des chercheurs ont commencé à adapter les drones de loisirs à des fins militaires. John Stuart Foster Jr., un physicien nucléaire qui a travaillé pour le gouvernement américain, était passionné par les modèles réduits d'avions et a envisagé l'idée d'y ajouter des armes. C'est ainsi que deux prototypes ont été développés en 1973 par l'agence de recherche de pointe du ministère de la Défense des États-Unis (DARPA), mais l'utilisation de drones similaires par Israël au cours de la dernière décennie, dans le cadre des conflits du Moyen-Orient, est ce qui a renouvelé l'intérêt pour les drones militaires et la poursuite de leur développement. Il est intéressant de remarquer que c'est en 1973 que des militaires ont abattu un drone pour la première fois, à l'aide d'un laser (lire l'article de *Popular Science* à l'adresse https://www.popsci.com/laser-guns-are-targeting-uavs-but-drones-are-fighting-back et celui de *Popular Mechanics* à l'adresse http://www.popularmechanics.com/military/research/a22627/drone-lasershotdown-1973/). Un drone a servi à tuer pour la première fois en 2001 en Afghanistan (voir https://www.theatlantic.com/international/archive/2015/05/america-first-drone-strike-afghanistan/394463/). Bien sûr, un opérateur humain se trouvait à l'autre extrémité de la détente.

La question de savoir s'il convient d'équiper les drones d'une IA fait l'objet de débats. Certains craignent que le processus de décision intégré aux drones puisse les conduire à faire des dégâts et à tuer des gens. Cependant, une IA pourrait aussi permettre aux drones

d’échapper plus facilement à la destruction ou d’exécuter d’autres tâches non destructives, de la même manière que des IA guident les voitures actuellement. Une IA pourrait même servir à stabiliser les mouvements du pilote dans des conditions météorologiques difficiles, un peu comme le système da Vinci aide les chirurgiens (voir la section « Assister un chirurgien » au Chapitre 7). Actuellement, les capacités meurtrières des drones militaires prêtent aussi à controverse parce que l’IA tendrait à faire d’un acte de guerre une chose abstraite et encore plus déshumanisée, qui se réduirait à des images transmises par les drones à leurs opérateurs et à des commandes émises à distance. Certes, ce serait toujours l’opérateur qui prendrait la décision de tuer, mais ce serait le drone qui exécuterait réellement la tâche, éloignant l’opérateur de la responsabilité de l’acte.



Ces considérations relatives aux drones militaires sont essentielles dans ce chapitre, car leur développement est lié à celui des drones civils et exerce une influence sur une grande partie des études actuelles relatives à cette technologie, par le biais de l’opinion publique. Par ailleurs, l’octroi d’une autonomie totale aux drones militaires inspire des histoires d’apocalypse sur fond d’IA qui dépassent le cadre de la science-fiction et deviennent une source de préoccupation du public. Pour une étude technique plus détaillée des modèles et de leurs capacités, lisez cet article de Deutsche Welle : http://www.dw.com/en/a-guide-to-military-drones/a-39441185.

Faire connaissance avec le quadrirotor

Souvent, les gens connaissent d’abord les drones à quatre rotors utilisés par les modélistes et autres passionnés, avant de découvrir les drones quadrirotors à usage

commercial (comme celui utilisé par Amazon dont il est question sur la page https://www.amazon.com/Amazon-Prime-Air/b?node=8037720011) dans le contexte de la révolution de la téléphonie mobile. Actuellement, la plupart des drones militaires ne sont pas de ce type, mais il en existe, par exemple le TIKAD de l'université Duke présenté à la page http://www.defenseone.com/tech-nology/2017/07/israeli-military-buyingcopter-drones-machine-guns/139199/ avec une démonstration sur la vidéo suivante : https://www.youtube.com/watch?v=VaTW8uAo_6s. Les drones de type hélicoptère ont d'abord été des prototypes pour les passionnés (voir http://www.popularmechanics.com/military/research/news/a27754/hobby-drone-sniper/).

Cependant, c'est grâce au téléphone mobile que tout cela a pu fonctionner. À mesure que les téléphones mobiles étaient de plus en plus petits, leurs batteries devenaient également plus petites et plus légères. Les téléphones mobiles intègrent aussi une caméra miniaturisée et une connectivité sans fil, autant de propriétés qu'un drone actuel doit posséder. Les petits drones d'il y a quelques décennies présentaient de nombreuses limitations :

» ils étaient commandés par radio à l'aide de consoles volumineuses ;

» il leur fallait une ligne de mire (sinon ils auraient volé de façon aveugle) ;

» c'étaient des miniaéronefs à ailes fixes (ils ne pouvaient donc pas faire du surplace) ;

» ils étaient propulsés par des moteurs diesel ou à essence bruyants (ce qui restreignait à la fois leur autonomie et la convivialité).

Depuis peu, des batteries légères de type lithium-polymère permettent aux drones :

- d'être équipés de moteurs électriques plus petits, plus silencieux et plus fiables ;
- d'être contrôlés à distance, par une communication non filaire ;
- de transmettre à l'opérateur des images vidéo (plus besoin d'une ligne de mire).

Aujourd'hui, les drones sont aussi équipés de GPS, d'accéléromètres et de gyroscopes, autant d'éléments qui ont aussi fait leur apparition dans la téléphonie mobile. Ces fonctionnalités permettent de contrôler la position, l'altitude et l'orientation, ce qui est utile également pour les applications de téléphonie mobile, mais qui est vraiment essentiel pour faire voler des drones.

Grâce à tous ces progrès, les drones, qui étaient initialement des copies miniatures d'avions à ailes fixes, sont devenus des sortes d'hélicoptères, sauf qu'ils ont besoin de plusieurs rotors pour pouvoir s'élever dans les airs et se diriger. L'utilisation de plusieurs rotors présente un avantage. Contrairement aux hélicoptères, les drones n'ont pas besoin de rotors à pas variable pour s'orienter. Ils utilisent des hélices simples à pas fixe qui, ensemble, reproduisent les mêmes fonctions que les rotors à pas variable, lesquels seraient plus coûteux et plus difficiles à contrôler. C'est la raison pour laquelle les drones que nous voyons aujourd'hui sont « multirotors » : ils peuvent être équipés de trois, quatre, six ou huit rotors. Parmi les différentes configurations possibles, le quadricoptère domine. C'est la configuration la plus répandue dans les utilisations commerciales et civiles. Il est équipé de quatre rotors (de petite taille), chacun orienté dans une direction, et l'utilisateur peut facilement le diriger et le faire tourner

en faisant varier l'orientation et la vitesse de chaque rotor (voir Figure 13.1).

Définir les usages des drones

Pour chaque type de drone, on distingue des applications actuelles et futures, et donc différentes possibilités d'utiliser l'IA. Les drones militaires de grande taille et de petite taille connaissent déjà une évolution parallèle en termes de technologie, et ils trouveront sans doute d'autres débouchés dans les domaines de la surveillance, du suivi et de l'action sur le terrain. Les spécialistes prévoient une extension des utilisations militaires aux drones à usage personnel et commercial, qui utilisent généralement une technologie différente (certains aspects sont communs, comme dans le cas du TIKAD de l'université Duke qui a commencé sa carrière dans le domaine des hobbies).

En dehors de l'utilisation dévoyée de drones de petite taille et facilement personnalisables par des insurgés ou des groupes terroristes (pour un exemple, consultez la page http://www.popularmechanics.com/military/weapons/a18577/isis-packing-drones-with-explosives/), les gouvernements s'intéressent de plus en plus aux petits drones pour les combats urbains et en intérieur. Dans les locaux fermés, pièces et couloirs, les capacités d'intervention de drones militaires comme Predator et Reaper, aussi grands que des avions, sont plutôt limitées (sauf s'il s'agit de détruire tout l'immeuble). Il en est de même de drones d'exploration et de surveillance comme les Raven et les Puma, qui sont faits pour des opérations sur le champ de bataille, et non pour les combats d'intérieur (une analyse détaillée de cette évolution militaire possible des drones grand public est présentée dans cet article de *Wired* :

https://www.wired.com/2017/01/military-may-soon-buy-drones-home/).



FIGURE 13.1 : Un quadricoptère vole en orientant ses rotors dans les directions qui conviennent.

Les drones à usage commercial sont loin de pouvoir être transférés directement sur le champ de bataille, même s'ils constituent la plateforme qui convient aux militaires pour développer différentes technologies. Une raison importante à l'utilisation des drones à usage commercial par les militaires est que les produits tout prêts sont considérablement moins onéreux que les armements habituels, si bien qu'ils sont à la fois facilement disponibles et utilisables par lots volumineux. Faciles à détourner et à modifier, ils nécessitent davantage de protection que les matériels militaires qui sont déjà protégés (leurs communications et leur contrôle peuvent être détournés par voie électronique), et il faut qu'y soient intégrés certains éléments matériels et logiciels avant qu'ils puissent être déployés pour une mission quelconque.

La navigation dans un espace fermé suppose des capacités renforcées pour pouvoir éviter les collisions,

recevoir des instructions sans GPS (les signaux de GPS ne sont pas facilement captés à l'intérieur d'un bâtiment) et engager le combat contre un ennemi potentiel. Par ailleurs, les drones auraient besoin de capacités de ciblage pour les missions de reconnaissance (détecter les embuscades et autres menaces) et pour atteindre eux-mêmes les cibles. De telles caractéristiques avancées ne sont pas présentes dans les technologies actuellement utilisées par les entreprises, et elles supposeraient une solution d'IA développée spécifiquement pour cela. Les chercheurs dans le domaine militaire s'emploient actuellement à développer les ajouts nécessaires pour obtenir un avantage militaire. De récentes avancées en matière de réseaux d'apprentissage profond réactifs installés sur un banal téléphone mobile, comme YOLO (https://pjreddie.com/darknet/yolo/) ou MobileNets, de Google (https://research.googleblog.com/2017/06/mobilenets-open-source-models-for.html), montrent comment l'intégration d'une IA avancée dans un petit drone est possible compte tenu des progrès actuels des technologies.

Voir des drones dans des rôles non militaires

Actuellement, les drones commerciaux n'ont pas beaucoup à offrir en regard des fonctionnalités avancées dont disposent les modèles militaires. Un drone commercial pourrait vous photographier avec ce qui vous entoure, à partir d'un point de vue aérien. Cependant, même avec les drones commerciaux, quelques utilisations innovantes vont se généraliser dans un avenir proche :

» la livraison de marchandises dans les temps, quelles que soient les conditions de circulation

(Google X, Amazon et de nombreuses startups travaillent à mettre au point cette nouveauté) ;

» le suivi pour la maintenance et la gestion de projet ;

» l'évaluation de différents types de dommages pour les assurances ;

» la cartographie et le comptage des cheptels pour les éleveurs ;

» l'assistance aux opérations de recherche et de sauvetage ;

» l'accès à l'Internet dans les zones isolées et non connectées (une idée en cours d'élaboration chez Facebook) ;

» la production d'électricité à partir des vents d'altitude ;

» le transport de personnes d'un point à un autre.

La possibilité de se faire livrer ses achats par un drone est une chose qui a rapidement attiré l'attention du public, grâce à une campagne de promotion entreprise par de grandes compagnies. Un des premiers innovateurs et des plus fameux est Amazon (qui promet que son service Amazon Prime Air sera bientôt opérationnel : https://www.amazon.com/Amazon-Prime-Air/b?node=8037720011). Google promet un service similaire avec son projet Wing (*https://www.wing.eu/fr*). Cependant, il se peut qu'il faille attendre encore des années avant de disposer d'un système de distribution par voie aérienne au moyen de drones, qui soit viable et évolutif.



Même si l'idée est de supprimer les intermédiaires dans la chaîne logistique pour des raisons de rentabilité, il reste à résoudre un certain nombre de problèmes techniques et d'ambiguïtés dans les réglementations. Les médias sont prompts à nous montrer des drones qui livrent des petits

colis, des pizzas ou des burritos dans des zones cibles à titre expérimental (https://www.theverge.com/2017/10/16/16486208/alphbetgoogle-project-wing-drone-delivery-testing-australia), mais en réalité, les drones ne peuvent pas aller loin ni transporter un poids important. Le plus gros problème est de réguler le vol de tous les drones qui devraient transporter quelque chose d'un point à un autre. Les drones posent des problèmes évidents : éviter des obstacles comme les lignes électriques, les bâtiments, et les autres drones, affronter de mauvaises conditions météorologiques, et trouver un endroit adapté pour atterrir à proximité immédiate du destinataire. Les drones doivent également éviter les espaces aériens sensibles, et satisfaire à toutes les exigences réglementaires imposées aux aéronefs. L'IA jouera un rôle clé dans la résolution d'un certain nombre de ces problèmes, mais pas de tous. Pour l'instant, il semblerait que les drones soient viables à petite échelle pour des activités de livraison plus importantes que la livraison à domicile de plats préparés : http://time.com/rwanda-drones-zipline/.

Les drones peuvent devenir vos yeux, c'est-à-dire assurer la vision dans des situations dans lesquelles voir par vous-même serait trop onéreux, trop dangereux ou trop difficile. Des drones commandés à distance ou semi-autonomes (utilisant des solutions d'IA pour la détection d'images ou le traitement des données mesurées par les capteurs) peuvent observer, surveiller, suivre ou chercher et secourir, sachant qu'ils peuvent voir du dessus n'importe quelle infrastructure et accompagner et assister des opérateurs humains sur demande dans leurs activités. Ainsi, des drones ont pu inspecter des lignes électriques, des pipelines et des infrastructures ferroviaires, ce qui permet le suivi d'infrastructures vitales mais difficilement accessibles de façon plus fréquente et moins coûteuse.

Même les compagnies d'assurance trouvent utile d'utiliser des drones pour évaluer les dommages.

Les drones sont utiles aux forces de police et aux premiers intervenants dans le monde entier pour des activités variées, des opérations de recherche et de sauvetage à la détection des incendies de forêt en passant par les missions de patrouille frontalière et la surveillance des foules. Les services de police trouvent de nouvelles manières d'utiliser les drones, notamment pour identifier les infractions au code de la route.

L'agriculture est aussi un domaine important dans lequel les drones révolutionnent les méthodes de travail. Non seulement ils peuvent surveiller les cultures et détecter les problèmes, mais ils peuvent servir à épandre les pesticides ou les engrais uniquement là où ils sont nécessaires, comme l'explique la revue *Technology Review* du MIT (https://www.technologyreview.com/s/526491/agricultural-drones/). Les drones permettent d'obtenir des images plus détaillées qu'un satellite, et à moindre coût, et ils peuvent être utilisés régulièrement pour assurer les fonctions suivantes :

» analyser les sols et cartographier le résultat grâce à l'analyse d'image et à des scanners laser 3D afin d'optimiser le rendement des semences et des plantations ;

» gérer les plantations en contrôlant les déplacements du tracteur ;

» suivre la croissance des cultures en temps réel ;

» épandre des produits chimiques où et quand il le faut ;

» irriguer où et quand il le faut ;

» évaluer la santé des cultures à l’aide d’une vision infrarouge, ce qu’un agriculteur ne peut pas faire.

L’*agriculture de précision* exploite les capacités de l’IA en termes de mouvement, de localisation, de vision et de détection. Cela peut permettre d’accroître la productivité (obtenir des cultures plus saines et davantage de nourriture pour la collectivité) tout en réduisant les coûts et les risques sanitaires lors des interventions humaines (plus besoin d’épandre des pesticides partout).

Les drones sont capables d’exploits plus admirables encore. Grâce aux drones, il est envisagé de déplacer les infrastructures existantes vers le ciel. Facebook, par exemple, cherche à fournir des connexions Internet (https://www.theguardian.com/technology/2017/jul/02/facebook-drone-aquila-internet-test-flight-arizona) là où les câbles de communication ne sont pas arrivés ou sont endommagés, en utilisant un modèle spécial, le drone Aquila (https://www.facebook.com/notes/mark-zuckerberg/the-technologybehind-aquila/10153916136506634/). Il existe aussi un projet d’utilisation de drones pour transporter les personnes et remplacer des moyens de transport courants comme la voiture (http://www.bbc.com/news/technology-41399406). Une autre possibilité consisterait à produire de l’électricité en altitude, là où les vents sont plus forts et où personne ne serait incommodé par le bruit des hélices.

Améliorer les drones grâce à l’IA

Que les drones soient utilisés comme jouets ou dans des applications commerciales ou militaires, l’IA est à la fois ce

qui rend la chose possible et ce qui change la donne. L'IA rend de nombreuses applications viables ou les améliore, en renforçant l'autonomie des systèmes et les capacités de coordination. Raffaello D'Andrea, un ingénieur italo-helvéto-canadien qui enseigne les systèmes dynamiques et les techniques de régulation à l'EPF de Zurich et qui invente des drones, fait une démonstration des avancées dans ce domaine dans cette vidéo : https://www.youtube.com/watch?v=RCXGpEmFbOw. La vidéo montre comment des drones peuvent devenir plus autonomes grâce aux algorithmes d'IA. L'autonomie change la manière dont le drone vole, elle réduit le rôle des humains, et permet la détection et la gestion automatiques des obstacles et une navigation sans risque dans les zones qui présentent des complications. La coordination implique la possibilité pour les drones de travailler ensemble sans avoir à rendre des comptes à une unité centrale ni à en recevoir des instructions, les drones devenant capables d'échanger des informations et de collaborer en temps réel à l'accomplissement d'une tâche donnée.

Si l'on pousse à l'extrême, l'autonomie peut même aller jusqu'à l'absence de toute personne guidant le drone. La machine volante détermine l'itinéraire à suivre et exécute d'elle-même des tâches spécifiques (l'être humain n'intervenant que pour les commandes de haut niveau). Les drones qui ne sont pas dirigés par un pilote utilisent un GPS pour optimiser leur itinéraire, mais cela n'est possible qu'en extérieur, et pas toujours avec une grande précision. En intérieur, le vol doit être plus précis, et cela suppose que le drone utilise davantage de données issues de ses capteurs pour caractériser les *abords de proximité* (les éléments d'un bâtiment, par exemple une protubérance sur un mur, avec laquelle il pourrait entrer en collision). Parmi ces capteurs, le moins onéreux et le plus léger est la caméra qui est intégrée à la plupart des

drones commerciaux. Cependant, la caméra ne suffit pas car il faut aussi un traitement performant des images à l'aide de la vision par ordinateur et de techniques d'apprentissage profond (un sujet abordé dans ce livre, par exemple au Chapitre 11 à propos des réseaux convolutifs).

Les entreprises anticipent l'exécution autonome de tâches par des drones commerciaux, qui seraient capables, par exemple, de livrer un colis de l'entrepôt au client final tout en gérant les problèmes éventuels tout au long du processus (comme dans le cas des robots, un problème peut toujours survenir, et l'appareil doit alors le résoudre sur place en utilisant l'IA). Au Jet Propulsion Laboratory, un laboratoire de la NASA à Pasadena, en Californie, des chercheurs ont récemment testé des drones dans le cadre d'un vol automatisé, en comparant la performance du dispositif à celle d'un pilote de drone professionnel qualifié (voir https://www.nasa.gov/feature/jpl/drone-race-human-versus-artificial-intelligence). De façon intéressante, le pilote humain a réalisé la meilleure performance au cours de ce test jusqu'au moment où il a ressenti de la fatigue, et c'est alors que les drones qui étaient plus lents, plus stables et moins sujets aux erreurs l'ont rattrapé. Dans l'avenir, on peut s'attendre au même phénomène que ce qui s'est produit avec les échecs et le Go : les drones automatisés feront mieux que les pilotes de drones, en termes de compétences de pilotage comme en termes d'endurance.

On peut aussi imaginer que grâce à des progrès dans la coordination, des centaines, sinon des milliers de drones pourront un jour voler ensemble. La capacité de coordination serait essentielle pour que les drones des particuliers et les drones commerciaux pullulent dans le ciel. Cette coordination serait appréciable pour éviter les collisions, pour partager les informations relatives aux obstacles, et pour analyser la circulation de façon similaire

à ce qui se pratique avec les voitures qui sont interconnectées et partiellement ou totalement automatisées (le Chapitre 14 est consacré aux voitures conduites par une IA).

La refonte des algorithmes qui commandent les drones est déjà en cours, et il existe dès à présent des solutions pour la coordination des activités des drones. Le MIT, par exemple, a récemment mis au point un algorithme de coordination décentralisée pour les drones (voir https://techcrunch.com/2016/04/22/mit-creates-a-control-algorithm-for-drone-swarms/). Cependant, la plupart des recherches ne sont pas connues, sachant que par nature, les possibilités d'utilisation de drones de façon coordonnée concernent avant tout l'armée. Des essaims de drones pourraient plus facilement pénétrer les défenses ennemies de façon discrète et effectuer des frappes difficiles à parer. Au lieu de devoir viser un seul drone, l'ennemi se retrouverait confronté à plusieurs centaines de petits drones qui envahiraient l'espace autour de lui. Il existe des solutions pour contrer des menaces similaires (voir http://www.popularmechanics.com/military/weapons/a23881/the-army-is-testing-a-real-life-phaser-weapon/). Un test récent sur un essaim de 100 drones (Perdix, un modèle sur mesure pour le ministère de la Défense des États-Unis) lâchés depuis trois avions F/A-18 Super Hornet pour effectuer des missions de reconnaissance et d'interception a été rendu public (https://www.technologyreview.com/s/603337/a-100-drone-swarm-dropped-from-jetsplans-its-own-moves/), mais d'autres pays se sont également lancés dans cette nouvelle course aux armements.

Quand l'entrepreneur Elon Musk, le cofondateur d'Apple Steve Wozniak, le physicien Stephen Hawking et bien



d'autres personnalités publiques et chercheurs en IA se sont inquiétés des récents progrès réalisés dans les armements assistés par l'IA, ils pensaient non pas à des robots comme ceux que l'on voit dans les films *Terminator* et *I, Robot*, mais à des drones armés et autres armes automatisées. Les armes autonomes pourraient bien être l'enjeu d'une nouvelle course aux armements qui changerait à jamais la conduite des guerres. Pour plus de détails sur ce sujet, consultez la page http://mashable.com/2017/08/20/ai-weapons-ban-open-letter-un/.

Prendre la mesure des problèmes de réglementation

Les drones ne sont évidemment pas les premiers ni les seuls objets à voler dans le ciel. Depuis des décennies, d'innombrables avions civils et militaires traversent les cieux chaque jour, et cela nécessite à la fois une réglementation stricte et un contrôle permanent pour garantir la sécurité des vols. Aux États-Unis, la Federal Aviation Administration (FAA) est l'organisme compétent pour réguler toute l'aviation civile et prendre des décisions concernant la gestion des aéroports et du trafic aérien. La FAA a publié une série de règles pour les systèmes d'aéronefs sans pilote (drones).

En août 2016, la FAA a publié une série de règles qui constituent ce qu'on a appelé Part 107. Ces règles portent sur l'utilisation commerciale des drones pendant la journée. La liste complète de ces règles est ici : https://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=20516. Elles sont au nombre de cinq :

- Voler au-dessous de 400 pieds (120 mètres) d'altitude.

- Voler à moins de 100 miles (161 km à l'heure).
- Garder en permanence l'aéronef dans son champ visuel.
- Détenir une licence adéquate.
- Ne jamais voler à proximité d'un aéronef piloté, en particulier à proximité d'un aéroport.
- Ne jamais survoler un attroupement, un stade ou un événement sportif.
- Ne jamais voler à proximité d'une intervention en situation d'urgence.

COMPRENDRE L'ORIENTATION DE L'APPRENTISSAGE

Ce livre est consacré pour une grande part à la création d'un environnement et à la fourniture de données permettant à une IA d'apprendre. Il examine aussi en long et en large ce qui est possible et ce qui ne l'est pas quand on utilise une IA, du point de vue de l'apprentissage. La morale et l'éthique sont même abordées, dans la mesure où elles concernent l'IA et ses utilisateurs humains. Cependant, l'orientation de l'apprentissage prévu pour une IA est également importante.

Dans le film *WarGames* (https://www.amazon.com/exec/obidos/ASIN/B0089J2818/datacservip0f-20/), l'ordinateur PROG (Plan de riposte opérationnelle de guerre) est doté d'une IA forte capable de déterminer la meilleure ligne de conduite en réponse à une menace. Dans la première partie du film, PROG cesse de se cantonner à un rôle de conseil et devient l'exécutant de la stratégie. C'est alors que survient un pirate qui veut jouer au jeu de la guerre thermonucléaire. Malheureusement, PROG ne fait pas la différence entre un jeu et la réalité, si bien qu'il lance réellement un plan d'engagement d'un conflit thermonucléaire avec l'Union soviétique. Ce film semble ne pas être loin de confirmer les pires craintes possibles concernant l'IA et la guerre.

Voici le plus étrange : le pirate, qui a été identifié et qui travaille maintenant pour les gentils, conçoit une méthode pour apprendre à l'IA la futilité. En d'autres termes, l'IA pénètre un environnement dans lequel elle apprend que gagner à certains jeux – en l'occurrence, le morpion – n'est pas possible. Quelle que soit la performance de chaque joueur, je jeu se termine toujours par une impasse. L'IA teste ensuite cette nouvelle connaissance dans le domaine de la guerre thermonucléaire. À la fin, l'IA conclut que la seule façon de gagner est de ne pas jouer du tout.

Dans la plupart des scénarios proposés par les médias, dans les romans de science-fiction et dans les films, l'environnement d'apprentissage n'est jamais évoqué. Pourtant, il s'agit d'un élément essentiel du problème, car ce que l'IA apprendra dépendra de la manière dont on aura configuré l'environnement. En matière d'équipement militaire, il est sans doute pertinent d'apprendre à l'IA à gagner, mais aussi de lui enseigner que dans certains scénarios gagner n'est pas envisageable, et que le meilleur coup consiste à ne pas jouer du tout.

La FAA publiera bientôt des règles relatives à l'utilisation des drones la nuit, précisant quand un drone peut ne plus être vu et dans quelles conditions il peut voler en milieu urbain, même s'il est actuellement possible d'obtenir de la FAA des dérogations. Le but de ces systèmes réglementaires est de protéger le public, sachant que l'impact des drones sur notre existence n'est pas encore bien connu. Ces règles permettent aussi à cette technologie de déboucher sur des innovations et sur une croissance économique.



En ce moment, tous les pays s'efforcent de réglementer l'utilisation des drones. Les réglementations servent à garantir la sécurité et favorisent l'utilisation des drones à des fins économiques. En France, par exemple, la loi autorise l'utilisation de drones dans l'agriculture avec un petit nombre de restrictions, ce qui en fait un pays pionnier en la matière.

Actuellement, compte tenu de l'absence d'IA, un drone peut facilement perdre sa connexion et se comporter de façon erratique, et parfois occasionner des dégâts (voir https://www.theatlantic.com/technology/archive/2017/03/drones-invisible-fence-president/518361/). Même si des mesures de sécurité sont prises pour certains modèles en cas de perte

de la connexion avec le responsable, par exemple le retour automatique du drone à l'endroit exact d'où il avait décollé, la FAA restreint leur utilisation au champ visuel de la personne qui les utilise.

Une autre mesure de sécurité importante est ce que l'on appelle le *gardiennage visuel* ou le *balisage géolocalisé*. Les drones qui utilisent le service GPS pour la localisation sont équipés d'un logiciel qui limite leur accès à des périmètres prédéterminés définis par des coordonnées GPS comme les aéroports, les zones militaires et autres domaines d'intérêt national. Vous pouvez obtenir une liste de paramètres à l'adresse http://tfr.faa.gov/tfr2/list.html. Pour en savoir davantage sur ce sujet, consultez la page https://www.theatlantic.com/technology/archive/2017/03/drones-invisiblefence-president/518361/.

Les algorithmes et l'IA viennent à la rescousse en préparant un environnement technologique approprié à l'utilisation sûre d'un ensemble de drones qui livreront des marchandises dans les grandes villes. L'Ames Research Center de la NASA travaille sur un système appelé *Unmanned Aerial Systems Traffic Management* (UTM) qui jouera le même rôle pour les drones qu'une tour de contrôle pour les avions. Toutefois, ce système est entièrement automatisé : son fonctionnement repose sur la capacité des drones à communiquer entre eux. UTM permettra d'identifier les drones dans le ciel (chaque drone sera défini par un identifiant, à l'image des plaques minéralogiques des voitures) et définira la trajectoire et l'altitude de croisière de chaque drone, afin d'éviter les risques de collisions, et de dommages pour les personnes. UTM devrait être mis à la disposition de la FAA pour une éventuelle mise en service ou pour des développements ultérieurs en 2019, ou un peu plus tard. Le site Internet de la NASA propose des informations supplémentaires sur ce

système de contrôle révolutionnaire qui pourrait rendre viable et sûre l'utilisation de drones commerciaux : https://utm.arc.nasa.gov/.



Si les restrictions ne suffisent pas et si certains drones malveillants représentent une menace, les forces de police et l'armée ont trouvé des contre-mesures efficaces : tirer sur le drone pour le faire tomber, l'attraper dans un filet, brouiller ses commandes, le faire tomber à l'aide d'un laser ou d'une émission de micro-ondes, ou même, tirer des missiles guidés.

Chapitre 14

Utiliser une voiture conduite par une IA

DANS CE CHAPITRE

- **Percevoir le progrès vers l'autonomie des véhicules**
- **Imaginer l'avenir avec des voitures qui se conduisent toutes seules**
- **Comprendre le cycle perception-planification-action**
- **Découvrir, utiliser et associer différents capteurs**

Une voiture sans conducteur est un *véhicule autonome* qui est capable de circuler depuis un point de départ jusqu'à une destination sans intervention humaine. L'autonomie ne consiste pas simplement à automatiser certaines tâches (comme dans la démonstration d'Active Park Assist sur la vidéo https://www.youtube.com/watch?v=xW-MhoLImqg), mais aussi à être capable de réaliser les étapes appropriées pour atteindre les objectifs de manière indépendante. Une voiture sans conducteur exécute d'elle-même toutes les tâches requises, une personne humaine étant éventuellement présente pour observer (et ne rien faire d'autre). Sachant que les voitures sans conducteur font partie de l'histoire depuis plus de 100 ans (oui, aussi incroyable que cela puisse paraître), ce chapitre commence par un bref rappel historique.



Pour connaître le succès, une technologie doit apporter un avantage qui soit considéré comme nécessaire et difficile à obtenir en utilisant d'autres méthodes. C'est pour cela que les voitures sans conducteur sont une chose aussi passionnante. La conduite mise à part, elles sont intéressantes à plus d'un titre. La section suivante de ce chapitre vous explique comment elles vont révolutionner la mobilité et vous permet de comprendre pourquoi cette technologie est si captivante.

Quand les voitures sans conducteur seront une chose un peu plus courante et quand elles seront admises comme faisant partie du quotidien, elles continueront à exercer un impact sur la société. La suite de ce chapitre vous aide à comprendre ces questions et leur importance. Vous y trouverez la réponse à la question de savoir ce que cela fera, de monter dans une voiture sans conducteur et de se dire que cette voiture va vous conduire sans problème d'un endroit à un autre.

Enfin, une voiture sans conducteur doit être équipée d'un certain nombre de types de capteurs pour pouvoir exécuter sa tâche. En effet, d'une certaine façon, on pourrait distinguer les capteurs qui voient de ceux qui entendent et de ceux qui touchent, mais ce serait là une simplification excessive. La dernière section de ce chapitre vous aide à comprendre le fonctionnement de ces divers capteurs et leur contribution au fonctionnement de la voiture sans chauffeur de façon globale.

Avoir un aperçu historique

Mettre au point des voitures qui se conduisent toutes seules fait partie depuis longtemps de la vision futuriste proposée par les récits et les films de science-fiction. Tout a commencé avec les premières expérimentations de voitures radiocommandées dans les années 1920. Pour

en savoir plus sur la longue et fascinante histoire des voitures autonomes, lisez cet article : https://qz.com/814019/driverless-cars-are-100-years-old/. Le problème de ces premiers véhicules était leur côté peu pratique : il fallait que quelqu'un les suive pour les guider au moyen d'une commande radio. Par conséquent, même si le rêve des voitures sans conducteur était caressé depuis si longtemps, les projets actuels ont peu de choses en commun avec ceux du passé, en dehors de la vision de l'autonomie.

Les voitures sans conducteur d'aujourd'hui sont profondément liées à des projets qui ont été lancés dans les années 1980 (https://www.technologyreview.com/s/602822/in-the-1980s-theself-driving-van-was-born/). Dans ces efforts plus récents, l'IA est mise à contribution pour éliminer le besoin d'une commande radio qui caractérisait les projets plus anciens. Ces projets sont financés par un certain nombre d'universités ainsi que par le secteur de la défense (plus particulièrement l'armée américaine). À un moment donné, l'objectif était de gagner le concours Grand Challenge de la DARPA qui a pris fin en 2007. Cependant, les préoccupations dans les domaines militaire et commercial fournissent aujourd'hui aux ingénieurs et aux développeurs de nombreuses raisons de continuer d'aller de l'avant.

Le tournant majeur dans ce concours a été la création du Stanley, conçu par le chercheur et entrepreneur Sebastian Thrun et son équipe. Ils ont remporté le Grand Challenge de la DARPA en 2005 (voir la vidéo à l'adresse https://www.youtube.com/watch?v=LZ3bbHTsOL4). Après cette victoire, Thrun a lancé le développement de voitures sans conducteur chez Google. Aujourd'hui, on peut voir le Stanley exposé au Musée

national d’histoire américaine, qui fait partie de la Smithsonian Institution.



Les militaires ne sont pas les seuls à stimuler la recherche dans le domaine des véhicules autonomes. Pendant longtemps, l’industrie automobile a souffert de surproduction, sachant qu’elle peut produire davantage de voitures que ce qu’exige la demande du marché. Or, la demande s’est réduite par suite de toutes sortes de pressions, entre autres la longévité des véhicules. Dans les années 1930, la longévité moyenne des automobiles était de 6,75 ans, tandis qu’elle est aujourd’hui égale ou supérieure à 10,8 ans et les voitures actuelles peuvent parcourir plus de 400000 km. La diminution des ventes a conduit certains constructeurs à sortir du secteur ou à fusionner pour former des compagnies plus grosses. La voiture sans conducteur est le remède miracle pour cette industrie, un moyen de remodeler favorablement la demande du marché et de convaincre les consommateurs de se moderniser. Cette technologie nécessaire débouchera sur un accroissement de la production, avec un grand nombre de nouveaux véhicules.

Avoir une vision claire de l’avenir de la mobilité

Si la voiture sans conducteur est une technologie de rupture, ce n’est pas simplement parce qu’elle change radicalement notre conception de l’automobile, c’est aussi parce que son apparition aura un impact significatif sur la société, sur l’économie et sur l’urbanisation. Pour le moment, il n’y a encore aucune voiture sans conducteur sur les routes. Il n’existe actuellement que des prototypes. Consultez par exemple cet article, qui parle de projets pilotes : https://www.wired.com/story/uber-self-driving-cars-pittsburgh/. Selon de

nombreuses sources, il faudra attendre encore une décennie pour que de telles voitures commencent à circuler, et le remplacement de tous les véhicules en circulation par des véhicules sans conducteur demandera bien plus de temps encore. Cependant, même si cette innovation fait encore partie de l'avenir, on peut en attendre beaucoup, comme le montrent les sections qui suivent.

Franchir les six niveaux d'autonomie

Il n'est pas possible de prédire la tournure que prendront les choses, mais nombreux sont ceux qui ont déjà spéculé sur les caractéristiques des voitures sans conducteurs. Pour plus de clarté, SAE International (http://www.sae.org/), un organisme de normalisation dans le domaine de l'automobile, a publié une norme de classification pour les véhicules autonomes (voir la norme J3016 sur la page https://www.smmt.co.uk/wp-content/uploads/sites/2/automated_driving.pdf). L'existence d'une norme permet de jalonner le progrès de l'automatisation des voitures. La norme SAE spécifie cinq niveaux d'autonomie :

» **Niveau 1 – assistance au conducteur :** le contrôle du véhicule est encore assuré par le conducteur, mais le véhicule peut prendre en charge des activités de soutien simples comme le contrôle de la vitesse. Ce niveau d'automatisation comprend le régulateur de vitesse, le contrôle de stabilité, et le freinage d'urgence assisté.

» **Niveau 2 – automatisation partielle :** la voiture peut plus souvent remplacer le conducteur, sachant qu'elle peut accélérer, freiner et changer de direction

si nécessaire. La responsabilité du conducteur est de rester attentif et de garder le contrôle de la voiture. Un exemple d'automatisation partielle est le freinage automatique, sur certains modèles, lorsque le véhicule détecte un risque immédiat de collision sur sa trajectoire (un piéton qui traverse, ou la voiture de devant qui s'arrête soudainement). On peut citer aussi le régulateur de vitesse intelligent (qui ne fait pas que stabiliser la vitesse du véhicule, mais qui adapte aussi la vitesse à la situation, par exemple si un autre véhicule se trouve devant vous), et le centrage du véhicule sur la voie. Ce niveau d'automatisation existe sur des modèles commercialisés depuis 2013.

» **Niveau 3 – automatisation conditionnelle :** la plupart des constructeurs automobiles planchent sur ce niveau d'automatisation au moment où ces lignes sont écrites. Le principe de l'automatisation conditionnelle est que le véhicule se conduit tout seul dans certaines circonstances (par exemple, uniquement sur autoroute, ou sur une route unidirectionnelle), au-dessous des limites de vitesse et sous le contrôle vigilant d'une personne humaine. Le système peut demander à la personne de reprendre le contrôle du véhicule. Un exemple de ce niveau d'automatisation est celui de certains modèles récents qui se conduisent tout seuls sur autoroute, et qui freinent automatiquement quand le trafic ralentit en raison d'un encombrement.

» **Niveau 4 – automatisme élevé :** le système du véhicule prend en charge toutes les tâches de la conduite (direction, accélération et freinage) et surveille et prend en compte tous les changements intervenant dans les conditions du trajet, du point de départ à la destination. L'intervention humaine n'est plus nécessaire, mais ce niveau d'automatisation

n’est possible qu’en certains endroits et dans certaines situations, si bien que le conducteur doit être en mesure de reprendre le contrôle du véhicule le cas échéant. Les constructeurs pensent pouvoir inaugurer ce niveau d’automatisation vers 2020.

» **Niveau 5 – automatisme intégral :** le véhicule circule du point de départ à la destination sans intervention humaine, avec un niveau de compétence comparable ou supérieur à celui d’un conducteur humain. Les voitures à automatisme intégral n’auront plus de volant. Ce niveau d’automatisation devrait être atteint vers 2025.

Même lorsque le cinquième niveau d’autonomie sera atteint, on ne verra pas ces véhicules automatiques circuler sur toutes les routes. Ils n’apparaîtront que dans un avenir éloigné, et il se pourrait bien qu’il y ait des difficultés à venir. La section « Pallier l’incertitude des perceptions », plus loin dans ce chapitre, parle des obstacles auxquels l’IA se heurtera lorsqu’il s’agira de conduire une voiture. La voiture sans conducteur n’est pas pour demain : elle sera probablement l’aboutissement d’une mutation progressive, avec l’apparition graduelle de modèles de plus en plus automatisés. Nous tiendrons le volant encore longtemps. Ce que l’on peut espérer voir bientôt apparaître, c’est une IA qui assistera le conducteur aussi bien au cours de ses trajets ordinaires que dans des conditions dangereuses, afin de rendre le déplacement plus sûr. Même lorsque des voitures sans conducteur seront commercialisées, le remplacement du parc existant demandera sans doute des années. En milieu urbain, cette révolution pourrait bien se faire attendre encore 30 ans.



Cette section mentionne un certain nombre de dates, et certains auront vite fait de croire qu’il s’agit d’échéances précises. Or, toutes sortes de facteurs peuvent accélérer

ou retarder l’adoption de la voiture sans conducteur. Les assureurs, par exemple, sont actuellement méfiants. Ils craignent que leurs produits d’assurance automobile ne puissent plus se vendre le jour où le risque d’accident sera devenu plus réduit (selon les prédictions de la société de conseil McKinsey, la fréquence des accidents sera réduite de 90 %.

Un lobbying des assureurs pourrait retarder l’adoption de la voiture sans conducteur. D’un autre côté, quiconque a déjà souffert de la perte d’un être cher dans un accident de voiture a des chances d’être favorable à tout ce qui pourrait réduire la fréquence des accidents. C’est là un facteur susceptible de faciliter l’adoption de cette nouveauté. Sachant que les pressions sociétales peuvent changer l’histoire de différentes manières, il n’est pas possible de prévoir une date précise.

Repenser le rôle de l’automobile dans notre existence

La mobilité est inextricablement liée à la civilisation. Il ne s’agit pas seulement du transport de personnes et des marchandises, il s’agit aussi de la circulation des idées sur de longues distances. Le jour où des automobiles ont commencé à circuler, rares étaient ceux qui pensaient qu’elles allaient bientôt remplacer les chevaux et les véhicules attelés. Pourtant, les automobiles présentent de sérieux avantages par rapport aux chevaux : leur entretien est plus pratique, elles vont plus vite, et elles peuvent parcourir de plus longues distances. Elles nécessitent aussi davantage de contrôle et d’attention, sachant que les chevaux sont conscients du chemin et réagissent quand des obstacles se présentent ou en cas de risque de

collision, mais ces exigences ne nous rebutent pas car l'avantage d'une plus grande mobilité n'est pas négligeable.

Aujourd'hui, le tissu urbain et l'économie sont très liés à l'automobile. C'est la voiture qui permet chaque jour à beaucoup de gens de parcourir de longues distances entre leur domicile et leur lieu de travail (ce qui permet le développement de l'immobilier dans les banlieues). Elle permet aussi aux entreprises de vendre facilement des produits à distance. L'automobile est un facteur de création d'entreprises et d'emplois, et les salariés de l'industrie automobile sont depuis longtemps les principaux acteurs d'une nouvelle redistribution des richesses. L'automobile est le premier vrai produit de marché de masse, fabriqué par des travailleurs pour d'autres travailleurs. Quand le marché de l'automobile est florissant, les collectivités qui lui sont liées le sont également ; quand il dépérit, une catastrophe peut s'ensuivre. Les trains et les avions ne peuvent effectuer que des trajets prédéterminés, contrairement aux voitures. L'automobile a apporté et étendu la mobilité sur une grande échelle, et elle a révolutionné le quotidien des gens davantage que les autres moyens de transport sur longue distance. Comme le disait Henry Ford, le fondateur de la Ford Motor Company : « la voiture a libéré les gens ordinaires des limitations de leur géographie. »

Comme lors de l'apparition de l'automobile, la civilisation est à la veille d'une nouvelle révolution amenée par la voiture sans conducteur. Le jour où l'autonomie de niveau 5 et la voiture sans conducteur se généraliseront, on peut s'attendre à des changements significatifs dans la conception des grandes villes et des banlieues, au sein de l'économie et dans le quotidien des gens. Certains changements coulent de source, d'autres sont moins faciles à entrevoir. Les changements les plus évidents et dont on parle le plus sont les suivants :

- **Moins d’accidents** : il y aura moins d’accidents car l’IA respectera le code de la route et tiendra compte des conditions de circulation : elle sera un conducteur plus intelligent que ne le sont les humains. La diminution du nombre d’accidents aura un impact profond sur la construction des voitures, qui offrent aujourd’hui davantage de sécurité que dans le passé, grâce aux protections structurelles passives. Dans l’avenir, compte tenu de la sécurité absolue qu’elles offriront, les voitures sans conducteur pourront être plus légères car elles auront besoin de moins de protections. Elles pourront même être en plastique. Elles consommeront donc moins de ressources. En outre, moins d’accidents signifie moins de coûts d’assurance, avec d’importantes conséquences pour le secteur des assurances concerné par l’économie des accidents.
- **Moins d’emplois de conducteurs :** de nombreux emplois de conducteurs disparaîtront. Ailleurs, le besoin de main-d’œuvre diminuera. Les coûts des transports de personnes et de marchandises en seront diminués d’autant, ce qui rendra ces transports plus accessibles. Le problème sera notamment de trouver de nouveaux emplois (rien qu’aux États-Unis, on estime que 3 millions de personnes ont un emploi dans les transports).
- **Plus de temps disponible :** l’automobile sans conducteur fera économiser du temps, une des ressources les plus précieuses. Elle ne fera pas voyager plus loin, mais les gens pourront employer autrement le temps qu’ils auraient dû consacrer à la conduite (puisque celle-ci sera assurée par l’IA). Par ailleurs, même si le trafic s’intensifie (en raison de la baisse des coûts de transport et d’autres facteurs), il sera plus fluide, et il n’y aura plus beaucoup

> d'encombrements. En outre, la capacité de transport des routes existantes augmentera. Cela peut sembler paradoxal, mais c'est là le pouvoir d'une IA lorsque l'humain n'intervient plus, comme l'illustre cette vidéo : https://www.youtube.com/watch?v=iHzzSao6ypE.

À ces effets immédiats s'ajoutent des implications subtiles que personne ne peut cerner tout de suite, mais qui peuvent apparaître évidentes après réflexion. Benedict Evans en souligne quelques-unes dans un article publié sur son blog, intitulé « Cars and second order conséquences » (http://ben-evans.com/benedictevans/2017/3/20/cars-and-secondorder-consequences). Dans cet article instructif, l'auteur observe plus attentivement les conséquences de l'apparition sur le marché aussi bien des voitures électriques que des voitures sans conducteur et du cinquième niveau d'autonomie. À titre d'exemple, les voitures sans conducteur pourraient faire du dystopique panoptique une réalité. Le panoptique est le bâtiment institutionnel théorisé par le philosophe anglais Jeremy Bentham à la fin du XVIIIe siècle, dans lequel tout le monde est sous surveillance sans le savoir. Quand des voitures sans conducteur circuleront en grand nombre, il pourra y avoir partout des voitures équipées de caméras, qui pourront tout enregistrer ou signaler. Votre propre voiture vous espionnera peut-être, ainsi que ceux qui vous entourent, au moment où vous vous y attendrez le moins.

Envisager l'avenir n'est pas un exercice facile, car il ne s'agit pas d'une simple question de causalité. Même l'examen d'un ordre des effets plus éloigné risque de se révéler vain si le contexte n'est plus celui que l'on avait anticipé. Il se peut, par exemple, qu'il n'y ait jamais de panoptique parce que le système juridique interdira la communication par les véhicules sans conducteur des

images filmées. C’est pourquoi les prévisionnistes se fondent sur des scénarios qui sont des descriptions approximatives d’un avenir possible : ces scénarios sont ou ne sont pas susceptibles de se réaliser, en fonction de diverses circonstances. Les spécialistes supposent que pour une auto dotée de capacités de conduite autonome, quatre scénarios différents peuvent être envisagés, qui redéfinissent chacun l’utilisation ou même la possession d’une voiture :

- **La conduite autonome sur de longs trajets sur autoroute :** si le conducteur peut décider de laisser l’IA le conduire à destination avec ses passagers, cela lui permet de consacrer son attention à d’autres activités. Ce premier scénario est souvent considéré comme une étape de lancement possible. Cependant, compte tenu de la vitesse de circulation élevée sur les autoroutes, abandonner le contrôle à une IA n’est pas entièrement dénué de risque, une collision pouvant toujours être provoquée par un autre véhicule guidé par une personne humaine. Il importe de considérer des conséquences comme les lois actuelles relatives à l’inattention au volant qui existent un peu partout. La question est de savoir si, dans le dispositif juridique, un conducteur qui recourt à une IA sera considéré comme inattentif. Il s’agit clairement d’un scénario d’autonomie de niveau 3.
- **Le service d’un voiturier** : dans ce scénario, l’IA intervient une fois que les passagers sont sortis de la voiture et leur évite le tracas de trouver une place pour stationner. Ils gagnent du temps, et l’automatisation débouche sur la possibilité à la fois d’une optimisation du stationnement (la voiture sans conducteur trouve l’emplacement le plus approprié) et d’un partage du véhicule (une fois que vous l’avez quitté, quelqu’un d’autre peut l’utiliser ; par la suite, vous trouverez vous-même une autre voiture à

proximité, sur le parking). Compte tenu des limitations de la conduite autonome, qui ne sert ici que pour le stationnement et la récupération du véhicule, ce scénario correspond à une transition entre le niveau 3 et le niveau 4.

- **Servir de chauffeur pour tout trajet, sauf là où les voitures sans conducteur restent interdites :** dans ce scénario avancé, l'IA conduit partout sauf là où elle est interdite pour des raisons de sécurité (par exemple sur les nouvelles infrastructures routières qui ne sont pas incluses dans la cartographie utilisée par le véhicule). Ici, la voiture sans conducteur est proche de sa maturité (autonomie de niveau 4).
- **Faire le taxi à la demande :** il s'agit d'une extension du deuxième scénario, les voitures sans conducteur ayant atteint une maturité suffisante pour pouvoir se conduire toutes seules en permanence (autonome de niveau 5), avec ou sans passagers, et assurer un service de transport pour quiconque le requiert. Dans ce scénario, les voitures font l'objet d'une utilisation intensive (alors qu'actuellement, elles sont stationnées 95 % du temps ; voir http://fortune.com/2016/03/13/cars-parked-95-percent-of-time/), ce qui révolutionne le concept de propriété d'un véhicule dans la mesure où l'on n'a plus besoin de posséder sa propre voiture.

Monter dans une voiture sans conducteur

La création d'une voiture sans conducteur, contrairement à ce que les gens s'imaginent, ne consiste pas à placer un robot sur le siège du conducteur et à lui faire conduire la

voiture. Pour conduire un véhicule, l’être humain exécute d’innombrables tâches qu’un robot ne saurait pas exécuter. La création d’une intelligence similaire à celle d’un être humain ferait appel à un certain nombre de systèmes reliés les uns aux autres et qui travailleraient ensemble, de façon harmonieuse, à définir un environnement de conduite approprié et sûr. Des efforts sont en cours pour obtenir une solution de bout en bout, plutôt que d’envisager une solution d’IA distincte pour chaque besoin. Le problème que pose la mise au point d’une voiture sans conducteur est qu’il s’agit de résoudre un grand nombre de petits problèmes et de faire en sorte que toutes les solutions fonctionnent bien ensemble. Ainsi, la reconnaissance de la signalisation et le changement de file, par exemple, supposent des systèmes distincts.

LES VOITURES SANS CONDUCTEUR ET LE DILEMME DU TRAMWAY

On entend dire que le problème de l'assurance et le dilemme du tramway risquent de gravement compromettre l'utilisation de voitures sans conducteur. Le problème de l'assurance consiste à se demander qui est responsable quand quelque chose tourne mal. Actuellement, des accidents se produisent, et les voitures sans conducteur devraient provoquer moins d'accidents que les humains, donc le problème semble pouvoir être facilement résolu par les constructeurs automobiles si elles ne sont pas assurées (les assureurs envisagent avec circonspection l'avènement des voitures automatisées car cela pourrait entraîner un remodelage important de leur activité principale). Des constructeurs qui se sont lancés dans ce genre de projet, comme Audi, Volvo, Google et Mercedes-Benz, se sont déjà engagés à accepter d'endosser la responsabilité des accidents qui seraient provoqués par leurs véhicules (voir http://cohen-lawyers.com/wp-content/uploads/2016/08/WestLaw-Automotive-Cohen-Commentary.pdf). Cela signifie que les constructeurs automobiles deviendront eux-mêmes les assureurs pour faciliter la mise sur le marché de voitures sans conducteur.

Le *dilemme du tramway* est un défi moral proposé par la philosophe britannique Philippa Foot en 1967 (mais c'est en réalité un dilemme bien plus ancien). Dans ce problème, un tramway fou est sur le point de tuer plusieurs personnes qui se trouvent sur la voie, mais il est possible de les sauver en aiguillant le tramway sur une autre voie, sauf que, dans ce cas, une autre personne sera tuée à leur place. Naturellement, il s'agit de choisir dans quelle direction il faut diriger le tramway, sachant que quelqu'un va mourir. Il existe un certain nombre de variantes du dilemme du tramway, et il existe même un site Internet du Massachusetts Institute of Technology (MIT), http://moralmachine.mit.edu/, qui propose des

alternatives plus adaptées aux situations qui pourraient se produire avec une voiture sans conducteur.

L'idée est qu'il peut se produire des situations dans lesquelles quelqu'un mourra, quelles que soient les capacités de l'IA qui conduit la voiture. Dans certains cas, le choix n'est pas entre deux inconnus, mais entre le conducteur et une personne qui se trouve sur la trajectoire. De telles situations se produisent même à l'heure actuelle, et nous résolvons le problème en laissant le choix moral entre les mains de la personne qui tient le volant. Certains sauveront leur propre personne, d'autres se sacrifieront pour sauver autrui, et certains choisiront ce qu'ils considéreront comme le moindre des deux maux ou le plus grand bien. La plupart du temps, ce sera une réaction instinctive face à la pression d'une menace et face à la peur. Mercedes-Benz, le plus ancien constructeur automobile du monde, a déclaré qu'il donnerait la priorité à la vie des passagers. Les constructeurs automobiles pourraient considérer qu'une situation catastrophique comme celle du dilemme du tramway est déjà très rare – et qu'elle deviendra plus rare encore avec les voitures sans conducteur – et que ce choix se justifie parce que nous avons nettement tendance, de façon innée, à vouloir nous protéger nous-mêmes.



On entend souvent parler de *solution de bout en bout* dans le domaine de l'IA, à propos du rôle de l'apprentissage profond. Étant donné le pouvoir de l'apprentissage par des exemples, beaucoup de problèmes ne nécessitent pas des solutions séparées. Il s'agit essentiellement d'une combinaison de plusieurs problèmes de moindre ampleur, à chacun correspondant une solution d'IA différente. L'apprentissage profond peut résoudre le problème de façon globale, en résolvant des exemples et en produisant une solution unique englobant tous les problèmes qui nécessitaient des solutions d'IA séparées dans le passé. Le problème est que pour le moment, l'apprentissage profond rencontre des limites

dans sa capacité d'y parvenir réellement. Une unique solution en termes d'apprentissage profond peut être viable pour certains problèmes, mais pour d'autres, il reste nécessaire de combiner des solutions d'IA de moindre ampleur pour pouvoir obtenir une solution globale qui soit fiable.

NVIDIA, qui produit des processeurs graphiques pour l'apprentissage profond, travaille actuellement à mettre au point des solutions de bout en bout. Regardez la vidéo https://www.youtube.com/watch?v=-96BEoXJMs0, qui montre l'efficacité d'une solution fondée sur l'exemple. Cependant, comme pour toute application d'apprentissage profond, la valeur de la solution dépend, dans une très large mesure, de l'exhaustivité et du nombre d'exemples utilisés. Pour que la voiture sans conducteur fonctionne comme une solution d'apprentissage profond de bout en bout, il faut un ensemble de données à partir duquel le système apprendra à conduire dans un nombre colossal de contextes et de situations. Un tel ensemble de données n'est pas encore disponible, mais il pourrait l'être un jour.



Néanmoins, on peut espérer que des solutions de bout en bout simplifieront la structure de la conduite autonome. L'article de la page https://devblogs.nvidia.com/parallelforall/explaining-deep-learning-self-driving-car/ explique comment fonctionne le processus d'apprentissage profond. Vous pouvez aussi lire l'article original de NVIDIA qui explique comment l'apprentissage de bout en bout permet de conduire une voiture : https://arxiv.org/pdf/1704.07911.pdf.

Réunir tous les aspects techniques

Sous le capot d'une voiture sans conducteur, il y a des systèmes qui fonctionnent ensemble selon le paradigme de la robotique : détection, planification et action. Tout commence au niveau de la détection, différents capteurs fournissant à la voiture des éléments d'information :

» Le GPS indique où se trouve le véhicule dans le monde (grâce à un système de cartographie), sous forme de coordonnées de latitude, de longitude et d'altitude.

» Le radar, les ultrasons et la télédétection par laser (lidar) détectent les objets et fournissent des données sur leur localisation et leurs mouvements sous forme de variations de leurs coordonnées dans l'espace.

» Les caméras informent la voiture sur ce qui l'entoure en lui fournissant des images sous un format numérique.



Une voiture sans conducteur est munie d'un certain nombre de capteurs spécialisés. La section « Pallier l'incertitude des perceptions », plus loin dans ce chapitre, les présente en détail et révèle la manière dont le système combine les informations qu'ils délivrent. Le système doit combiner et traiter les données provenant des capteurs pour que les perceptions nécessaires au fonctionnement de la voiture deviennent utiles. La combinaison des données provenant des capteurs définit donc différentes perceptions de l'environnement dans lequel se trouve le véhicule.

La *localisation* consiste à déterminer où la voiture se trouve dans le monde, une tâche exécutée principalement en traitant les données provenant du GPS. Le GPS est un système de navigation par satellite initialement créé à des fins militaires. Utilisé à des fins civiles, il présente une certaine inadéquation (si bien que seul un personnel

autorisé peut l'utiliser avec toute la précision qu'il offre). On retrouve le même genre d'inadéquation dans d'autres systèmes comme GLONASS (le système de navigation russe), GALILEO (ou GNSS, le système européen) et le BeiDou (ou BDS, le système chinois). En conséquence, quelle que soit la constellation de satellites utilisée, la voiture peut déterminer qu'elle est sur une certaine route, mais elle peut ne pas savoir sur quelle file elle roule (ou même finir par rouler sur une route parallèle). Outre la localisation imprécise que fournit le GPS, le système traite les données du GPS conjointement avec les données du capteur à laser pour déterminer la position exacte d'après les détails de l'environnement immédiat.

Le *système de détection* détermine ce qui est autour de la voiture. Ce système fait appel à un certain nombre de sous-systèmes qui accomplissent chacun une tâche spécifique en utilisant une série unique de données de capteurs et en effectuant une analyse de ces données :

» La détection des files se fait grâce à l'analyse des images d'une caméra ou à la segmentation des images par des réseaux spécialisés d'apprentissage profond, l'image étant divisée en plusieurs zones selon leur type (route, voitures, et piétons).

» La détection et la classification des panneaux et des feux de signalisation sont réalisées au moyen d'un traitement des images provenant des caméras, grâce à des réseaux d'apprentissage profond qui déterminent tout d'abord la zone de l'image dans laquelle se trouve le panneau ou le feu, puis en leur affectant le type qui convient (le type de signalisation ou la couleur des feux).

» Les données combinées du radar, du laser, des ultrasons et des caméras permettent de localiser les objets extérieurs et de suivre leurs mouvements en termes de direction, de vitesse et d'accélération.

» Les données du laser (lidar) sont principalement utilisées pour détecter l'espace libre sur la route (une voie dégagée ou un emplacement de stationnement libre).

Faire entrer en scène l'IA

Après la phase de détection, consistant à permettre à la voiture sans conducteur de déterminer où elle se trouve et quels événements se produisent autour d'elle, la phase de planification commence. Dès ce moment, l'IA joue un rôle clé. La planification consiste à exécuter les tâches suivantes :

» **Route** : déterminer la trajectoire à suivre. Sachant que vous voyagez en vue de vous rendre à un endroit précis (ce n'est pas toujours vrai, mais c'est vrai la plupart du temps), vous souhaitez atteindre votre destination de la manière la plus rapide et la plus sûre. Dans certains cas, le coût doit aussi être pris en compte. Les algorithmes de routage, qui sont des algorithmes classiques, sont utiles ici.

» **Prédiction de l'environnement** : permettre à la voiture de se projeter dans l'avenir, sachant qu'il faut du temps pour prendre la mesure d'une situation, décider d'une manœuvre, et l'exécuter. Durant le temps nécessaire à la manœuvre, d'autres voitures peuvent décider de changer de position, ou effectuer également des manœuvres. En conduisant, on s'efforce aussi de déterminer ce que les autres conducteurs sont sur le point de faire, afin d'éviter des collisions possibles. Une voiture sans conducteur fait la même chose, grâce à la prédiction par l'apprentissage machine. Elle estime ce qui va se produire et tient compte du futur.

- **Planification comportementale** : l'intelligence de la voiture sans conducteur est essentiellement ici. Il s'agit des activités qui sont nécessaires pour que la voiture reste sur la route : rester sur la même voie, changer de voie, converger ou déboucher sur une route, conserver une distance de sécurité, obéir à la signalisation lumineuse, aux stops et aux panneaux prescrivant de céder le passage, éviter les obstacles, et bien d'autres choses encore. Toutes ces tâches sont exécutées grâce à l'IA, qui est comme un système expert intégrant le savoir-faire d'un certain nombre de conducteurs ou un modèle probabiliste comme le réseau bayésien, ou même un modèle d'apprentissage machine, plus simple.
- **Planification de trajectoire :** déterminer la manière dont la voiture va exécuter les tâches requises, sachant qu'il existe généralement plus d'une façon d'atteindre un objectif. Si la voiture décide de changer de voie, par exemple, il ne faut pas qu'elle accélère brutalement, ni qu'elle se rapproche trop d'un autre véhicule. Il faut que la manœuvre soit acceptable, sans risque, et agréable.

Se rendre compte qu'il ne s'agit pas simplement de l'IA

Après la détection et la planification, le moment est venu d'agir. La détection, la planification et l'action sont les différentes parties d'un cycle qui se répète jusqu'à ce que la voiture ait atteint sa destination, se gare et s'immobilise. Les actions consistent à accélérer, à freiner, et à orienter la direction. Les instructions sont décidées au cours de la phase de planification, et la voiture ne fait qu'exécuter les actions avec l'aide d'un système de contrôle comme la

commande proportionnelle, intégrale et dérivée (PID) ou le contrôleur à modèle prédictif (MPC), des algorithmes qui vérifient si les actions prescrites s'exécutent correctement et, sinon, prescrivent immédiatement des contre-mesures adéquates.

Cela peut sembler un peu compliqué, mais il s'agit simplement de trois systèmes qui interviennent l'un après l'autre, du point de départ à la destination. Chaque système est constitué de sous-systèmes qui résolvent chacun un seul problème de conduite, comme le décrit la Figure 14.1, à l'aide des algorithmes les plus rapides et les plus fiables.

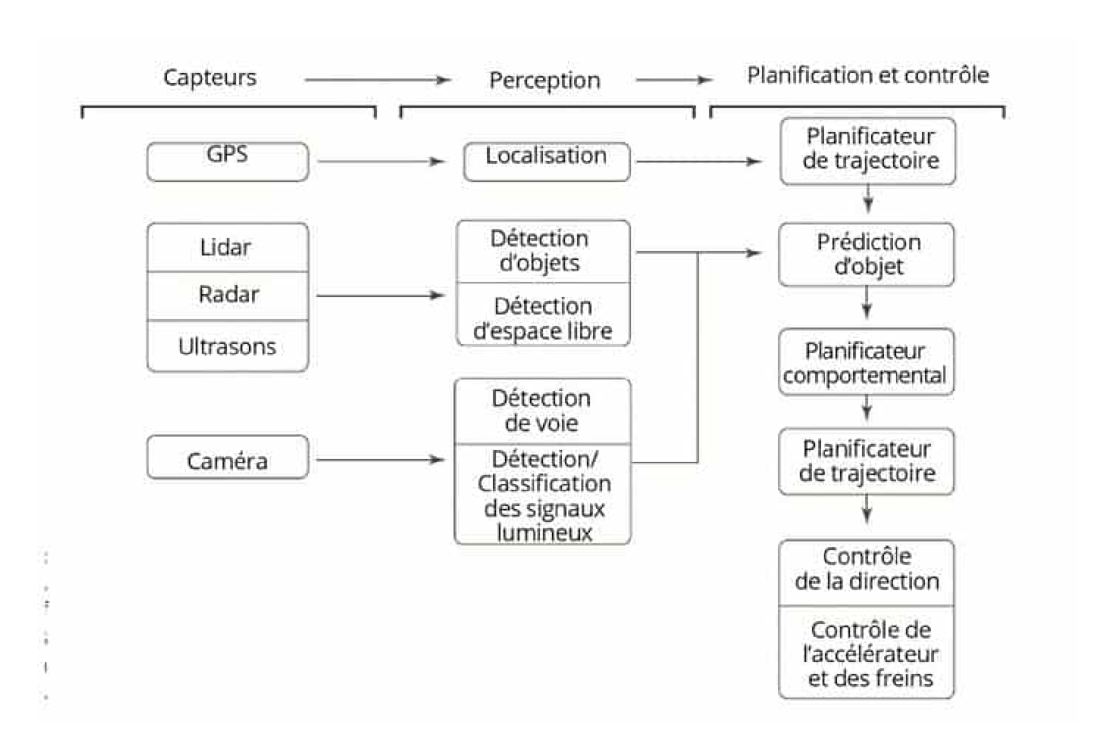


FIGURE 14.1 : Une représentation globale et schématique des systèmes qui interviennent dans une voiture sans conducteur.

Au moment où ces lignes sont écrites, ce cadre représente la pointe du progrès. Les voitures sans conducteur continueront sans doute d'être l'association de systèmes matériels et logiciels assurant différentes fonctions et opérations. Dans certains cas, les systèmes

présenteront des fonctionnalités redondantes, par exemple si plusieurs capteurs détectent le même objet extérieur, ou s'il est fait appel à plusieurs systèmes de traitement de la perception pour vérifier que vous vous trouvez sur la bonne file. La redondance permet l'absence d'erreurs, et donc une diminution du nombre de morts. Ainsi, même si un système comme un détecteur de signalisation fondé sur l'apprentissage profond ne fonctionne pas bien ou est mis en défaut, d'autres systèmes peuvent en assurer la fonction et minimiser ou éviter entièrement les conséquences pour la voiture.

Pallier l'incertitude des perceptions

Steven Pinker, professeur au Département de psychologie de l'université de Harvard, explique dans son livre *L'Instinct du langage* que « dans la robotique, les problèmes faciles sont difficiles et les problèmes difficiles sont faciles ». En effet, une IA est incroyablement performante quand il s'agit de jouer aux échecs contre un champion, mais des activités plus triviales comme saisir un objet sur une table, éviter une collision avec un piéton, reconnaître un visage ou répondre correctement à une question au téléphone, peuvent se révéler très difficiles à réussir pour une IA.



Le *Paradoxe de Moravec* dit que ce qui est facile pour les humains est difficile pour l'IA (et inversement), comme l'ont expliqué dans les années 1980 les spécialistes en robotique et sciences cognitives Hans Moravec, Rodney Brooks et Marvin Minsk. Les humains ont eu amplement le temps de développer des aptitudes comme marcher, courir, saisir un objet, parler et voir : ces aptitudes se sont développées à travers l'évolution et la sélection naturelle, pendant des millions d'années. Pour pouvoir survivre dans

ce monde, les humains font ce que tous les êtres vivants ont fait depuis que la vie existe sur terre. Au contraire, l'abstraction et les mathématiques sont une découverte relativement récente pour nous, et nous n'y sommes pas naturellement adaptés.

Les voitures présentent certains avantages par rapport aux robots, qui doivent évoluer à l'intérieur des bâtiments et en extérieur. Elles se déplacent sur des voies et des surfaces créées spécifiquement pour elles et qui sont généralement bien cartographiées, et elles bénéficient déjà de solutions mécaniques éprouvées pour cela.

Les actionneurs ne sont pas le problème le plus important pour les voitures sans conducteur. Ce qui présente de sérieuses difficultés, ce sont la planification et la détection. La planification se fait à un niveau élevé (pour lequel l'IA excelle généralement). En ce qui concerne la planification générale, les voitures sans conducteur peuvent déjà compter sur les navigateurs GPS, un type d'IA spécialisé dans la fourniture d'orientations. La détection est le véritable goulet d'étranglement pour les voitures sans conducteur, car sans elle, il n'y a pas de planification ni de manœuvre possible. Un conducteur détecte la route en permanence pour pouvoir maintenir son véhicule sur sa voie, faire attention aux obstacles et respecter les règles qui s'imposent.



Le matériel de détection est actualisé continuellement à ce stade de l'évolution des voitures sans conducteur afin d'obtenir des solutions plus fiables, plus précises et moins onéreuses. Par ailleurs, le traitement et l'exploitation des données provenant des capteurs dépendent d'algorithmes performants, comme le *filtre de Kalman* qui existe déjà depuis plusieurs décennies.

Avoir un aperçu des capteurs sur un véhicule

Les capteurs sont les éléments essentiels à la perception de l'environnement, et une voiture sans conducteur dispose de deux types de détection, interne et externe :

- **Les capteurs proprioceptifs :** ils détectent l'état du véhicule, c'est-à-dire la situation des systèmes (moteur, transmission, freins et direction) et la position du véhicule dans le monde grâce à la localisation par le GPS, la rotation des roues, la vitesse du véhicule et son accélération.
- **Les capteurs extéroceptifs :** ce sont les caméras, les systèmes à laser (lidar), les radars et les capteurs à ultrasons, qui détectent l'environnement immédiat.

Les capteurs proprioceptifs et extéroceptifs contribuent ensemble à l'autonomie du véhicule. La localisation par le GPS, en particulier, indique (éventuellement sous forme d'estimation grossière) où se trouve le véhicule, ce qui est utile à haut niveau pour planifier les directions à prendre, et les actions permettant de conduire la voiture à sa destination. Le GPS guide la voiture sans conducteur sur sa trajectoire comme il guide un conducteur, en fournissant les bonnes indications.

Les capteurs extéroceptifs (Figure 14.2) guident plus spécifiquement la conduite du véhicule. Ils remplacent ou renforcent la perception humaine dans une situation donnée. Chaque capteur perçoit l'environnement d'une manière différente, chaque capteur présente des limitations, et chaque capteur excelle dans son propre domaine de perception.

Les limitations prennent différentes formes. Dans l'étude de la fonction exercée par les capteurs, il convient de tenir compte du coût, de la sensibilité à la lumière, de la sensibilité aux conditions météorologiques, du bruit enregistré (la sensibilité du capteur change, avec des conséquences en termes de précision), du champ du capteur et de la résolution. Par ailleurs, les capteurs peuvent indiquer avec une grande précision la vitesse, la position, la hauteur ou la distance des objets, et ils permettent aussi de connaître la nature des objets et de savoir comment les classer.

La caméra

Les caméras sont des capteurs passifs qui produisent une vision. Cette vision peut être ou non stéréoscopique. Compte tenu de leur faible coût, il est possible d'en disposer un grand nombre sur le pare-brise ainsi que sur la calandre, sur les rétroviseurs latéraux, sur le hayon et sur la lunette arrière. Les caméras à vision stéréoscopique imitent généralement la perception humaine et captent l'information à partir de la route et des véhicules environnants, tandis que les caméras à vision simple servent généralement à détecter les panneaux de signalisation et la signalisation lumineuse. Les données qu'elles transmettent sont traitées par des algorithmes de traitement d'image ou par des réseaux neuronaux d'apprentissage profond de manière à obtenir des informations en termes de détection et de classification (par exemple, l'identification d'un feu rouge ou d'un panneau de limitation de vitesse). Les caméras peuvent être à haute résolution (elles sont alors capables de détecter des petits détails), mais elles sont sensibles à la lumière et aux conditions météorologiques (obscurité, brouillard et neige).

Le lidar (LIght Detection And Ranging)

Le lidar utilise des rayons infrarouges (d’une longueur d’onde voisine de 900 nanomètres, ils sont invisibles à l’œil humain) pour estimer la distance entre le capteur et l’objet atteint. Un système de rotation permet de projeter le rayon sur une certaine amplitude et de fournir les estimations sous forme d’un nuage de points de collision, ce qui permet d’évaluer les formes et les distances. Selon le prix (les plus coûteux étant généralement les meilleurs), ces dispositifs peuvent offrir une résolution plus élevée que les radars. Ils sont cependant plus fragiles et se salissent plus facilement, car ils sont placés à l’extérieur du véhicule (le lidar est le dispositif pivotant que l’on voit au-dessus de la voiture de Google dans ce reportage de CBS : https://www.youtube.com/watch?v=_qE5VzuYFPU).

Le radar (RAdio Detection And Ranging)

Le radar émet des ondes radio qui rebondissent sur la cible qu’elles rencontrent, et la durée de leur trajet définit la distance et la vitesse. Les radars peuvent être disposés dans les pare-chocs avant et arrière ainsi que sur les côtés de la voiture. Les voitures particulières en sont déjà équipées depuis des années. Ils permettent d’utiliser un régulateur de vitesse intelligent, un avertisseur d’angles morts, un avertisseur de collision, et un système d’évitement. Contrairement aux autres capteurs avec lesquels un certain nombre de mesures successives sont nécessaires, un radar peut détecter la vitesse d’un objet à partir d’une impulsion unique, en raison de l’effet Doppler (voir
http://www.physicsclassroom.com/class/waves

/Lesson-3/The-Doppler-Effect). Il existe des versions à courte et longue portée, qui donnent l'une et l'autre une représentation de l'environnement immédiat et peuvent être utilisées à des fins de localisation. Le radar est moins sensible que les autres types de détecteurs aux conditions météorologiques, en particulier la pluie et le brouillard, il balaye sur un angle de vue de 150 degrés et sa portée varie entre 30 et 200 mètres. Ses principales faiblesses sont le manque de résolution (il ne transmet pas beaucoup de détails) et son incapacité à détecter correctement les objets statiques.

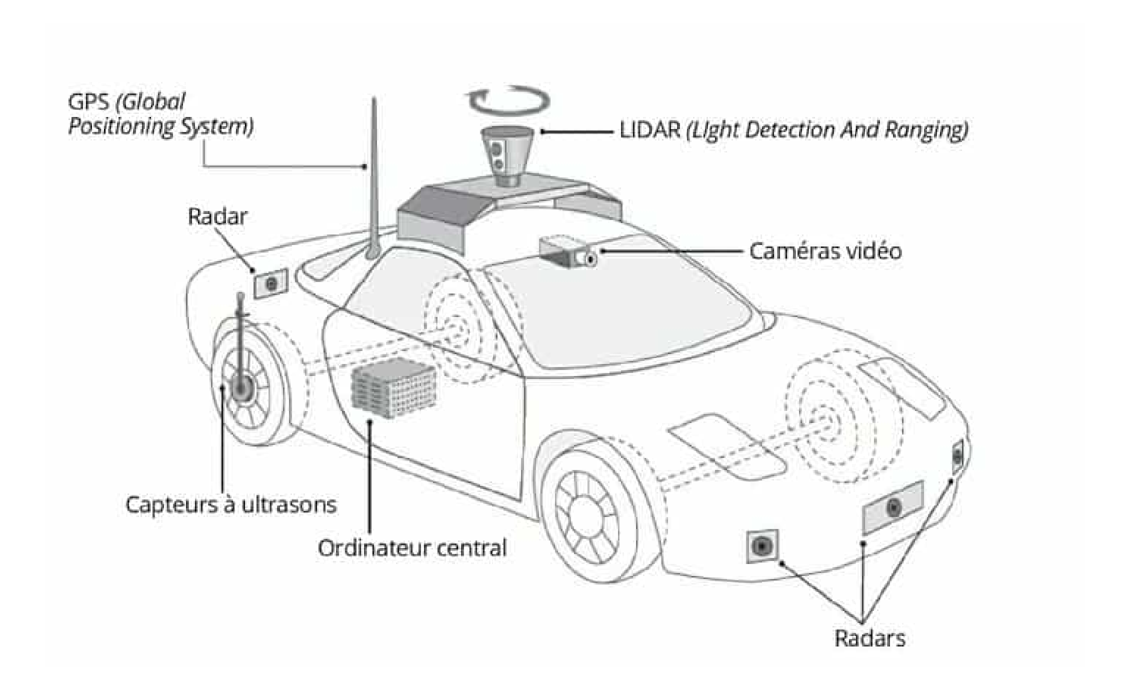


FIGURE 14.2 : Représentation schématique des capteurs extéroceptifs sur une voiture sans conducteur

Les capteurs à ultrasons

Les capteurs à ultrasons sont similaires aux radars, sauf qu'ils utilisent non pas des micro-ondes, mais des sons à haute fréquence (c'est-à-dire des ultrasons, inaudibles pour l'être humain mais audibles pour certains animaux). Le principal inconvénient des capteurs à ultrasons (utilisés

par les constructeurs à la place des lidars, qui sont plus fragiles et coûtent plus cher) est que leur portée est limitée.

Faire la synthèse de ce que l'on perçoit

Concernant la détection de l'environnement immédiat du véhicule, on utilise différentes mesures en fonction des capteurs installés. Or, chaque type de capteur offre des niveaux différents de résolution, d'amplitude et de sensibilité au bruit, si bien qu'on obtient des mesures différentes pour une même situation. En d'autres termes, aucun système n'est parfait, et les faiblesses des capteurs empêchent parfois que la détection soit correcte. Les signaux des sonars et des radars peuvent être absorbés, et les rayons du lidar peuvent traverser des solides transparents. En outre, les caméras peuvent être induites en erreur par des reflets ou un mauvais éclairage, comme l'explique cet article de la revue technique du MIT *Technology Review* : https://www.technologyreview.com/s/608321/this-image-is-why-self-driving-cars-come-loaded-with-many-types-ofsensors/.

Les voitures sans conducteur serviront à accroître notre mobilité, c'est-à-dire à préserver nos vies et celles des autres. On ne peut pas permettre qu'une voiture sans conducteur manque de détecter un piéton qui apparaîtrait soudainement devant sa calandre. Pour des raisons de sécurité, les constructeurs consacrent beaucoup d'efforts à la fusion des capteurs, qui consiste à combiner les données provenant des différents capteurs afin d'obtenir une mesure unifiée qui soit préférable à toute mesure simple. La fusion des capteurs résulte le plus souvent de l'utilisation de variantes du filtre de Kalman (comme le

filtre de Kalman étendu ou le filtre de Kalman « unscented », encore plus complexe). Rudolf E. Kálmán était un ingénieur électricien et inventeur hongrois, qui a émigré aux États-Unis pendant la Seconde Guerre mondiale. Pour son invention, qui a trouvé de nombreuses applications dans le guidage, la navigation et le contrôle des véhicules, aussi bien dans les voitures que dans les avions et les engins spatiaux, Kálmán a reçu en 2009 des mains du président Obama la médaille nationale des sciences.

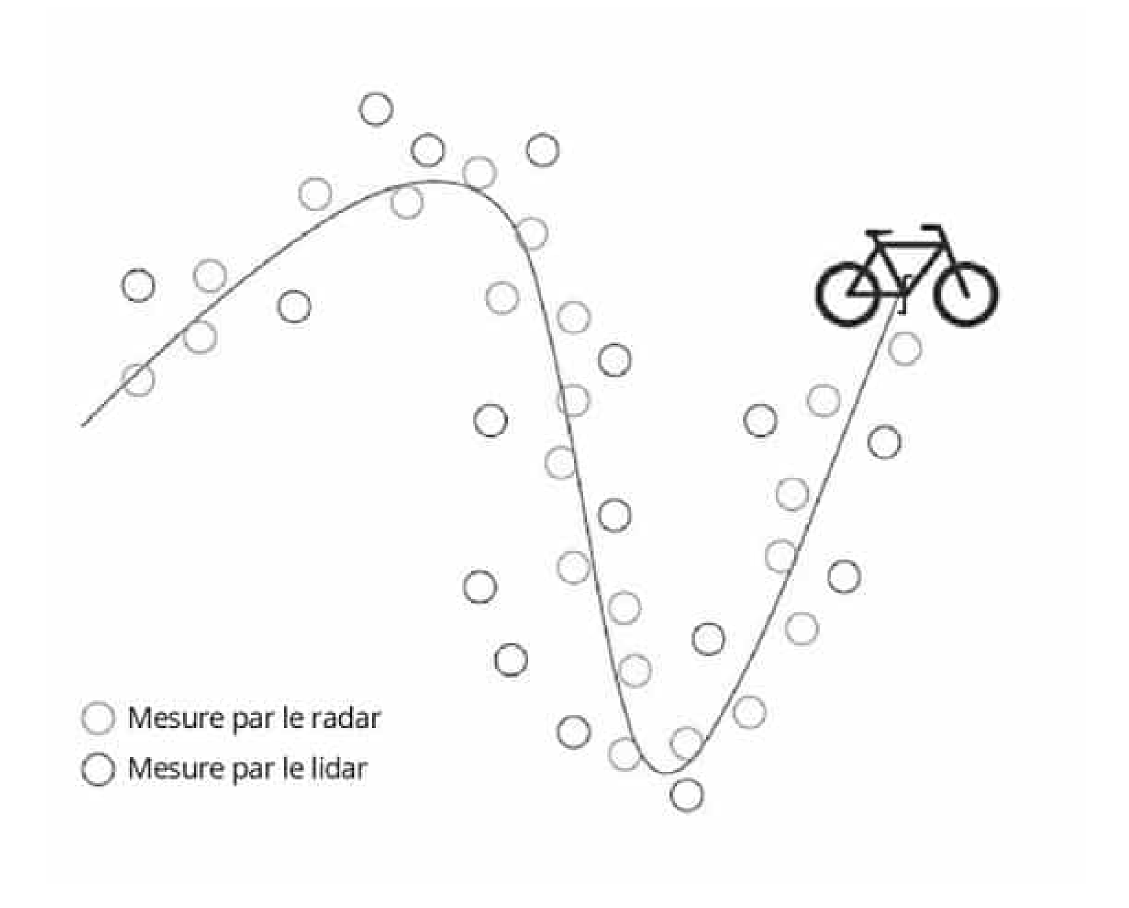


FIGURE 14.3 : Ce filtre de Kalman évalue la trajectoire d'une bicyclette en fusionnant les données d'un radar et d'un lidar.

Un algorithme de filtre de Kalman filtre les diverses observations effectuées dans le temps sous forme d'une séquence unique de mesures constituant une estimation

réelle (les mesures initiales étant des manifestations inexactes). Il commence par traiter l'ensemble des mesures d'un objet détecté (phase de prédiction d'état) pour estimer la position courante de l'objet. Ensuite, à mesure que de nouvelles observations arrivent, il utilise les nouveaux résultats qu'il obtient et met à jour les précédents, afin d'obtenir une estimation plus fiable de la position et de la vitesse de l'objet (phase d'actualisation des mesures), comme le montre la Figure 14.3.

De cette manière, une voiture sans conducteur peut alimenter l'algorithme avec les observations des capteurs afin d'obtenir une évaluation résultante des objets environnants. L'évaluation combine les points forts des différents capteurs tout en évitant leurs points faibles, ce qui est rendu possible parce que le filtre utilise une version plus élaborée des probabilités et du théorème de Bayes (voir Chapitre 10).

PARTIE 5
Se pencher sur l'avenir de l'IA

DANS CETTE PARTIE…

» Déterminer dans quel cas une application ne fonctionnera pas

» Envisager l'utilisation de l'IA dans l'espace

» Créer de nouvelles occupations humaines

Chapitre 15

Avoir un aperçu d'une application qui ne mène nulle part

DANS CE CHAPITRE

- » **Définir des scénarios d'utilisation de l'IA**
- » **Comprendre ce qui se produit quand l'IA échoue**
- » **Élaborer des solutions pour des problèmes qui n'existent pas**

Les chapitres précédents expliquaient ce qu'est et ce que n'est pas une IA, quels problèmes elle peut bien résoudre et quels problèmes sont apparemment hors des limites de ses applications. Même en disposant de toutes ces informations, vous pouvez facilement reconnaître un projet d'application qui ne verra même pas la lumière du jour parce que l'IA n'est pas en mesure de répondre au besoin concerné. Ce chapitre traite des applications qui ne mènent nulle part. Peut-être aurait-on pu l'intituler plutôt « Pourquoi on a encore besoin des humains », mais le titre choisi est plus parlant.

Dans ce chapitre, vous allez découvrir ce qui se produit quand on essaie de développer des applications qui ne mènent nulle part. La conséquence la plus préoccupante est l'hiver de l'IA. Un *hiver de l'IA* est ce qui se produit dès lors que les promesses des promoteurs de l'IA vont au-

delà de ce qu'ils sont en mesure d'offrir, si bien que les entreprises concernées perdent leurs financements.

L'IA peut aussi déboucher sur le piège que constitue l'élaboration de solutions à des problèmes qui n'existent pas réellement. En effet, on peut être émerveillé par les splendeurs d'une solution, mais si cette solution ne répond pas à un besoin réel, personne ne l'achètera. L'essor d'une technologie n'est possible que lorsque les utilisateurs sont prêts à dépenser de l'argent pour l'obtenir parce qu'elle répond à des besoins. Ce chapitre se termine par un aperçu des solutions à des problèmes qui n'existent pas.

Utiliser l'IA là où elle ne fonctionnera pas

Le Tableau 1.1 du Chapitre 1 recense les sept types d'intelligence. Dans une société entièrement fonctionnelle, on les retrouve tous les sept, et différentes personnes excellent dans différents types d'intelligence. En combinant les efforts de tout le monde, on peut exploiter ces sept types d'intelligence de manière à satisfaire les besoins de la société.



Au vu du Tableau 1.1, on a vite fait de remarquer que l'IA fait l'impasse totale sur deux types d'intelligence et ne présente que des capacités modestes par rapport à trois autres. L'IA excelle dans le domaine des mathématiques, de la logique et de l'intelligence kinesthésique, sa capacité se limitant à résoudre différentes sortes de problèmes qu'une société entièrement fonctionnelle a besoin de résoudre. Les sections qui suivent décrivent des situations dans lesquelles l'IA ne peut pas fonctionner pour la simple raison qu'il s'agit d'une technologie, et non d'une personne.

Définir les limites de l’IA

Quand vous parlez à Alexa, vous pouvez oublier que c’est à une machine que vous vous adressez. Or, la machine n’a aucune idée de ce que vous dites, elle ne vous perçoit pas comme une personne et elle n’a aucun réel désir d’avoir un échange avec vous : elle agit uniquement en fonction des algorithmes développés pour la faire fonctionner et des données que vous lui fournissez. Cela n’empêche pas que les résultats soient admirables. On a vite fait de verser dans l’anthropomorphisme sans s’en rendre compte, et de voir Alexa comme l’extension de quelque entité humanoïde. Cependant, il manque à l’IA les éléments essentiels dont il est question dans les sections suivantes.

La créativité

Vous trouverez une infinie variété d’articles, de sites, de musiques, d’œuvres d’art, d’écrits et de toutes sortes d’autres productions censément créatives issues d’une IA. Le problème est que l’IA ne peut rien créer. La créativité implique des manières de penser. Beethoven, par exemple, avait une manière particulière de penser la musique. On peut reconnaître qu’un morceau de musique classique est de Beethoven même sans connaître toute son œuvre, car sa musique possède certaines particularités qui sont liées à la manière dont pensait Beethoven.

Une IA peut créer un nouveau morceau de Beethoven grâce à une visualisation mathématique de son processus de pensée, au moyen d’un apprentissage basé sur des exemples de musiques de Beethoven. Le fondement résultant, pour la création d’un nouveau morceau de Beethoven, est de nature mathématique. C’est cette logique mathématique qui permet à une IA, par exemple,

de jouer du Beethoven comme le joueraient les Beatles : https://techcrunch.com/2016/04/29/paul-mccartificial-intelligence/.



Le problème, quand on veut faire de la créativité avec des mathématiques, c'est que les mathématiques ne sont pas créatives. Être créatif, c'est développer une nouvelle forme de pensée, quelque chose que personne n'avait encore jamais connu (pour plus de détails, voir https://www.csun.edu/~vcpsy00h/creativity/define.htm). La créativité ne consiste pas simplement à penser différemment : elle consiste à définir une nouvelle façon de penser.

La créativité suppose aussi le développement d'une perspective différente, ce qui revient essentiellement à définir un type différent de série de données (pour rester dans l'optique des mathématiques). Une IA se limite aux données qui lui sont fournies. Elle ne peut pas créer ses propres données, mais seulement des variations sur les données existantes qui sont celles à partir desquelles elle a réalisé son apprentissage. L'encadré « Comprendre l'orientation de l'apprentissage », au Chapitre 13, développe cette idée de perspective. Pour pouvoir apprendre à une IA quelque chose de nouveau, de différent et de surprenant, il faut entreprendre de lui fournir l'orientation appropriée.

L'imagination

Créer, c'est définir quelque chose de réel, qu'il s'agisse de musique, d'art, d'écriture ou de toute autre activité produisant un résultat qui puisse être vu, entendu, touché ou appréhendé de quelque manière par autrui. L'imagination est l'abstraction de la création, si bien qu'elle est plus éloignée encore des limites des capacités de l'IA. Une personne peut imaginer des choses qui ne sont pas réelles et qui ne pourront jamais être réelles.

L'imagination, c'est l'esprit qui vagabonde à travers les domaines d'activité et qui joue avec ce qui pourrait exister si les règles existantes n'y faisaient pas obstacle. La vraie créativité est souvent le résultat d'une imagination fructueuse.

D'un point de vue purement humain, tout le monde peut imaginer quelque chose. L'imagination nous différencie de tout le reste et nous met souvent dans des situations qui n'ont rien de réel. L'article du *Huffington Post* à l'adresse https://www.huffingtonpost.com/lamishaserfwalls/5-reasons-imagination-is-_b_6096368.html mentionne cinq raisons pour lesquelles l'imagination est essentielle si l'on veut dépasser les limites de la réalité.

De même qu'une IA ne peut pas créer de nouveaux modes de pensée ni élaborer de nouvelles données sans utiliser les sources existantes, elle doit aussi exister dans les limites de la réalité. En conséquence, il y a peu de chances qu'on parvienne un jour à mettre au point une IA qui serait douée d'imagination. L'imagination suppose non seulement une intelligence créative, mais aussi une intelligence intrapersonnelle, or une IA ne possède aucune de ces deux formes d'intelligence.



L'imagination, comme un certain nombre de caractéristiques humaines, est de nature émotionnelle. Une IA n'a pas d'émotions. En fait, si l'on veut savoir ce que peut faire une IA par rapport à ce que peut faire un être humain, il est souvent utile de se poser simplement la question de savoir si la tâche fait appel à l'émotion.

Des idées originales

Imaginer quelque chose, créer quelque chose de réel à partir de ce qui a été imaginé, puis exploiter cet exemple réel d'une chose qui n'avait jamais existé auparavant,

c’est développer une idée. Pour créer une idée avec succès, il faut une bonne intelligence créative, intrapersonnelle et interpersonnelle. Créer une nouveauté est formidable quand on veut définir des versions uniques de quelque chose ou se distraire. Cependant, pour que cette nouveauté devienne une idée, il faut la partager, de telle sorte que les autres la voient aussi.

Des déficiences dans les données

La section « Tenir compte des cinq types de données incorrectes » du Chapitre 2 parle des problèmes de données qu’une IA doit résoudre pour pouvoir exécuter les tâches pour lesquelles elle est conçue. Le seul problème est qu’une IA n’est généralement pas capable de reconnaître les données fausses, à moins de disposer d’une pléthore d’exemples de données ne comportant pas de telles erreurs, ce qui peut être plus difficile à obtenir que ce que vous pensez. Les humains, en revanche, décèlent souvent les erreurs avec une facilité relative. Parce qu’il a vu plus d’exemples qu’une IA ne pourra jamais en voir, un humain peut identifier les données fausses. Grâce à son imagination et à sa créativité, il peut se représenter ce qui est faux d’une manière qui est impossible pour une IA, car l’IA est ancrée dans la réalité.



Les possibilités d’informations fausses dans les données sont si variées qu’il n’est pas toujours possible d’en faire l’inventaire. Les humains en ajoutent souvent sans même y penser. Les éviter est parfois impossible, sachant qu’elles peuvent être liées à un point de vue, à un biais ou à un certain cadre de référence. Comme une IA ne peut pas les identifier toutes, les données servant à la prise de décisions présenteront toujours un certain degré de déficience. Le fait que cette déficience puisse affecter la capacité de l’IA à produire un résultat exploitable dépend

de son type et de son degré, ainsi que de la performance des algorithmes.

Cependant, le type le plus curieux de déficience dans les données à prendre en compte est l'information fausse que quelqu'un veut délibérément obtenir comme résultat. Cette situation se produit plus souvent qu'on le pense, et le seul moyen de remédier à ce problème particulier d'origine humaine réside dans la communication subtile que permet l'intelligence interpersonnelle qui fait défaut à l'IA. Supposons, par exemple, qu'une personne s'achète de nouveaux vêtements, et qu'ils soient très moches, du moins, à vos yeux (même si votre jugement sur une tenue vestimentaire peut être subjectif). Or, si vous êtes intelligent, vous allez la complimenter sur son choix. En effet, cette personne n'attend pas de vous un avis impartial, elle espère plutôt que vous allez la soutenir et l'approuver. La question que percevrait l'IA serait « Est-ce que ces vêtements me vont ? », mais la question réelle est « Est-ce que tu m'approuves ? » ou « Approuves-tu ma décision d'acheter ces vêtements ? » Vous pouvez remédier en partie au problème en lui conseillant des accessoires pour compléter sa tenue, ou bien lui faire comprendre subtilement qu'elle ne pourrait peut-être pas paraître en public dans cette tenue.

Il y a aussi le problème d'éviter la vérité qui blesse, qu'une IA ne pourra jamais résoudre parce qu'elle n'a pas d'émotion. Une *vérité blessante* est une vérité qui, au lieu de profiter à son destinataire, le fait souffrir, que cette souffrance soit émotionnelle, physique ou intellectuelle. Prenons le cas d'une personne qui ignore qu'un de ses parents, aujourd'hui décédé, avait été infidèle à l'autre. Sachant qu'ils ne sont plus là, cette information n'a plus aucune pertinence et le mieux serait que cette personne reste heureuse de l'ignorer. Supposons maintenant qu'elle entende quelqu'un raconter en détail cette infidélité. Elle n'y gagnera rien, et cela lui fera du mal. Une IA pourrait

causer un mal similaire en révélant cette vérité après avoir analysé une série de rapports de police, de registres d'hôtels, de reçus et d'autres sources. Toutefois, dans ce dernier cas, la vérité est révélée parce que l'IA n'a pas d'intelligence émotionnelle (empathie), si bien qu'elle est incapable de comprendre le besoin qu'a cette personne de cultiver une image idéalisée de ses parents. Malheureusement, même quand un jeu de données contient des informations assez fiables pour que l'IA produise un résultat exploitable, ce résultat peut être plus nuisible qu'utile.

Appliquer l'IA de façon incorrecte

Les limites de l'IA déterminent le domaine des possibilités de l'exploiter correctement. Or, même à l'intérieur de ce domaine de possibilités, on peut obtenir un résultat inattendu et indésirable. Ce peut être le cas si, après avoir fourni des données en entrée à une IA, on lui demande de calculer la probabilité de certains événements. Lorsqu'elle dispose de données suffisantes, l'IA peut produire un résultat mathématiquement cohérent avec ces données. Cependant, l'IA ne peut pas produire de nouvelles données, ni créer des solutions fondées sur ces données, ni imaginer un moyen de pallier un manque de données, ni trouver des idées pour mettre en œuvre une solution. Tout cela reste du domaine de l'humain. Tout ce que l'on peut attendre de l'IA, c'est une prédiction de probabilité.



Les résultats de l'IA sont souvent basés sur les probabilités ou sur les statistiques. Malheureusement, ces méthodes mathématiques ne sont pas applicables aux individus, mais seulement aux groupes. L'utilisation des statistiques génère d'innombrables problèmes dès qu'il ne s'agit pas d'obtenir un résultat concret, par exemple

lorsqu'il s'agit de conduire une voiture. L'article de la page http://public.wsu.edu/~taflinge/evistats.html traite des problèmes que pose l'utilisation des statistiques. Lorsqu'une application de l'IA représente un enjeu pour des individus, il faut parer à l'imprévu, et prévoir même l'impossibilité totale d'atteindre les objectifs que l'on s'était fixés.

Un autre problème réside dans la possibilité que les données à traiter reflètent une certaine opinion, ce qui est bien plus courant qu'on pourrait le penser. La différence entre une opinion et un fait est que le fait peut être entièrement prouvé, auquel cas tout le monde conviendra qu'il est véridique (sauf peut-être ceux qui ont un esprit très fermé) ; tandis que les opinions reflètent le manque ou l'absence de preuves scientifiques à l'appui des données. Par ailleurs, les opinions sont liées aux émotions. Même face aux preuves définitives du contraire, certains préfèrent s'en tenir à leur opinion plutôt qu'aux faits. Nos opinions nous confortent, pas les faits. Or, une IA échouera généralement lorsqu'une opinion est en jeu. Même avec le meilleur algorithme, le résultat ne contentera jamais tout le monde.

Entrer dans un monde d'attentes irréalistes

Les sections qui précèdent expliquent qu'attendre d'une IA qu'elle exécute certaines tâches ou l'utiliser dans des situations insuffisamment concrètes ne peut qu'engendrer des problèmes. Malheureusement, beaucoup de gens persistent à envisager de confier à l'IA des tâches qu'elle ne pourra jamais accomplir. Ces attentes irréalistes ont des causes variées :

» **Les médias :** les livres, les films, et plus généralement tous les médias visent à susciter en

nous une réponse émotionnelle. Or, cette réponse émotionnelle est précisément la source d'attentes irréalistes. Nous nous imaginons qu'une IA peut faire telle ou telle chose, alors que dans le monde réel, il est impossible qu'elle y parvienne.

- **L'anthropomorphisme** : outre les émotions suscitées par les médias, nous avons aussi tendance à nous attacher à des objets. Certains donnent un nom à leur voiture comme s'il s'agissait d'une personne, ils lui parlent, et peut-être se demandent-ils si elle ne se sent pas bien quand elle est en panne. Or, une IA ne peut ni sentir, ni comprendre, ni communiquer (en réalité), elle ne sait rien faire d'autre que traiter des nombres : énormément de nombres. Quand on attend de l'IA qu'elle manifeste soudain des sentiments et se comporte comme un être humain, l'entreprise est vouée à l'échec.
- **Un problème qui n'est pas défini :** une IA ne peut résoudre qu'un problème qui soit bien défini. En présentant à une personne une série d'entrées possibles, vous pouvez attendre d'elle qu'elle trouve une question correspondante à partir d'une extrapolation. Supposons qu'une série d'épreuves se traduise principalement par des échecs, mais que certaines personnes réussissent les épreuves et atteignent l'objectif désiré. On pourrait imaginer une IA qui améliorerait les résultats des épreuves grâce à une interpolation, en identifiant de nouveaux candidats dont les caractéristiques seraient celles des personnes qui ont réussi. Cependant, une personne humaine pourrait améliorer les résultats par extrapolation en cherchant pourquoi certains ont réussi et en identifiant la cause, que cette cause concerne les caractéristiques des sujets ou non (il se peut que les conditions environnementales aient

changé ou que le sujet ait simplement une attitude différente). Cependant, pour qu'une IA puisse résoudre un problème, il faut qu'un humain puisse exprimer ce problème d'une manière qui soit compréhensible pour l'IA. Les problèmes non définis, ceux qui sortent du cadre de l'expérience humaine, ne peuvent pas être résolus en utilisant une IA.

» **Une technologie déficiente :** souvent, dans ce livre, il est question d'un problème qui ne pouvait pas être résolu à une certaine époque, faute de disposer de la technologie adéquate. En effet, il n'est pas réaliste d'attendre d'une IA qu'elle résolve un problème alors que la technologie est insuffisante. Dans les années 1960, par exemple, faute de disposer des capteurs et de la puissance de traitement dont on dispose aujourd'hui, il aurait été impossible de produire une voiture sans conducteur. Depuis, les progrès réalisés dans ces technologies ont rendu ce genre de projet possible.

Mesurer les effets des hivers de l'IA

Les hivers de l'IA se produisent lorsque les chercheurs et autres promoteurs de l'IA ont promis des bienfaits qui ne se concrétisent pas dans le délai attendu, si bien que le financement de l'IA se tarit et que la recherche ne se poursuit qu'à pas de tortue. Depuis 1956, le monde a connu deux hivers de l'IA (actuellement, nous sommes dans le troisième été de l'IA). Les sections qui suivent examinent plus en détail les causes, les effets et les résultats de l'hiver de l'IA.

Comprendre l’hiver de l’IA

Il est difficile de dire précisément à quel moment l’IA a commencé. Les Grecs de l’Antiquité rêvaient déjà à la création d’hommes mécaniques, comme en témoignent les histoires d’Héphaïstos et de la Galatée de Pygmalion dans la mythologie grecque, et l’on peut supposer que ces créatures humanoïdes étaient censées posséder une certaine forme d’intelligence. On pourrait donc dire que le premier hiver de l’IA s’est produit entre la chute de l’Empire romain et le temps où des alchimistes ont imaginé insuffler un esprit dans la manière, au Moyen Âge, avec le Takwin de Jābir ibn Hayyān, l’homuncule de Paracelse et le Golem du rabbin Judah Loew. Cependant, il ne s’agissait que d’histoires sans fondement. Les premiers projets à caractère scientifique datent de 1956, avec le lancement des premières recherches subventionnées sur l’intelligence artificielle à Dartmouth College.



L’hiver de l’IA se produit quand le financement de l’IA décline. Le mot *hiver* est approprié, car de la même manière qu’un arbre continue sa croissance en hiver même si c’est à un rythme ralenti, comme en témoigne l’observation de ses anneaux de croissance, l’IA a continué à progresser, quoiqu’à un rythme glaciaire, au cours des hivers de l’IA entre 1974 et 1980 et à nouveau entre 1987 et 1993.

Définir les causes de l’hiver de l’IA

On peut considérer l’hiver de l’IA comme la conséquence de promesses farfelues et impossibles à tenir. Au démarrage des premiers projets à Dartmouth College, en 1956, les futures éminences de la recherche en IA

avaient prédit qu'il ne faudrait pas plus d'une génération pour qu'un ordinateur aussi intelligent qu'un être humain voie le jour. C'était il y a plus de soixante ans, et les ordinateurs sont encore loin d'approcher l'intelligence humaine. Si vous avez lu les chapitres précédents, vous savez que les ordinateurs ne seront sans doute jamais aussi intelligents que les humains, du moins en ce qui concerne certains types d'intelligence (et pour l'instant ils n'ont dépassé les capacités humaines que dans très peu de domaines).



Le problème que posent ces promesses exagérées, concernant les capacités de l'IA, est lié à la croyance qu'avaient ses premiers promoteurs à la possibilité de formaliser toute la pensée humaine sous forme d'algorithmes. En réalité, cette idée avait déjà germé chez les philosophes chinois, indiens et grecs. Cependant, comme le montre le Tableau 1.1, au Chapitre 1, seules certaines composantes de l'intelligence humaine peuvent être formalisées. Le meilleur résultat possible serait la mécanisation du raisonnement mathématique et logique, mais dans les années 1920 et 1930, David Hilbert avait mis les mathématiciens au défi de prouver que tout raisonnement mathématique pouvait être formalisé. La réponse à ce défi avait été donnée par la preuve d'incomplétude de Gödel, la machine de Turing et le lambda calcul de Church. Il en était ressorti d'une part qu'il n'était pas possible de formaliser n'importe quel raisonnement mathématique, et d'autre part, que dans les domaines où la formalisation était possible, on pouvait aussi mécaniser le raisonnement : c'est là le fondement de l'IA.

Une autre partie du problème des promesses exagérées est l'excès d'optimisme. Au cours des premières années de l'IA, les ordinateurs résolvaient des problèmes algébriques, prouvaient des théorèmes de géométrie et apprenaient à parler anglais. Les deux premiers résultats

sont raisonnables quand on considère que l'ordinateur ne fait qu'analyser les données entrantes et les mettre sous une forme qui lui permette de les manipuler. Le problème vient du troisième résultat. L'ordinateur ne parlait pas réellement en anglais : en réalité, il convertissait des données textuelles en schémas numériques, lesquels étaient à leur tour convertis en résultats analogiques, et cela ressemblait à de la parole, mais ce n'en était pas. L'ordinateur ne comprenait absolument rien à l'anglais, ni à aucune autre langue. Certes, les chercheurs entendaient de l'anglais, mais l'ordinateur ne faisait rien d'autre que traiter des zéros et des un sous des formats spécifiques, sans les considérer du tout comme du langage.



Même les chercheurs ont souvent eu tendance à surestimer les capacités de l'ordinateur. Ainsi, ELIZA, créée par Joseph Weizenbaum, semblait capable d'entendre ce qu'on lui disait et de répondre de façon intelligente. Malheureusement, les réponses étaient préparées et mises en boîte et l'application, en réalité, n'entendait rien, ne comprenait rien et ne disait rien. ELIZA a pourtant été le premier agent conversationnel et elle a bien représenté un pas en avant, quoique minuscule. Simplement, le battage qui a été fait était tout à fait disproportionné. C'est le problème que l'IA rencontre aujourd'hui. Les gens sont déçus quand ils se rendent compte qu'il n'y a rien de réel derrière ce qu'on leur a vanté. Les scientifiques et les promoteurs de l'IA continuent à courir à l'échec en mettant en avant le tape-à-l'œil plutôt que la technologie réelle. Le premier hiver de l'IA a été provoqué par ce genre de prédictions :

> » **H.A. Simon** : « D'ici une dizaine d'années, un ordinateur numérique sera champion mondial d'échecs » (1958) et « Dans vingt ans, les machines seront capables de faire n'importe quel travail que fait un être humain. » (1965)

> » **Allen Newell** : « D’ici une dizaine d’années, un ordinateur numérique aura découvert et prouvé un nouveau théorème mathématique important. » (1958)
>
> » **Marvin Minsky** : « Dans une génération (…) le problème que pose la création d’une ‘intelligence artificielle’ sera résolu pour l’essentiel » (1967) et « D’ici trois à huit ans, nous aurons une machine possédant l’intelligence générale d’un être humain. » (1970)

Aujourd’hui, au vu de ces prétentions farfelues, on n’a pas de mal à comprendre pourquoi les gouvernements ont cessé de financer les recherches. À propos des arguments exprimés à l’encontre de ces prédictions, y compris par des spécialistes de l’IA, consultez la section « Analyser l’argument de la chambre chinoise » du Chapitre 5.

Le second hiver de l’IA est arrivé pour les mêmes raisons que le premier : promesses exagérées, surexcitation et optimisme excessif. Cette fois, l’essor de l’IA a commencé avec les systèmes experts, une sorte de programme d’IA qui résout des problèmes en utilisant des règles logiques. Par ailleurs, les Japonais sont entrés dans la course avec leur projet Ordinateurs de cinquième génération, un système informatique basé sur une architecture de traitement parallèle massif. L’idée était de mettre au point un ordinateur capable d’exécuter beaucoup de tâches en parallèle, de façon similaire au cerveau humain. Finalement, John Hopfield et David Rumelhart ont fait revivre le connexionnisme, une stratégie consistant à modéliser les processus mentaux sous forme de réseaux constitués d’unités simples et interconnectés.

Cela s’est terminé comme une sorte de bulle économique. Les systèmes experts se sont révélés fragiles, même

quand ils tournaient sur des systèmes informatiques spécialisés. Les systèmes en question sont finalement apparus comme des gouffres économiques facilement remplaçables par de nouveaux systèmes informatiques courants, moyennant un coût significativement réduit. Le projet Ordinateurs de cinquième génération des Japonais a aussi été victime de ce phénomène de bulle : le système s'est révélé extrêmement coûteux à fabriquer et à entretenir.

Réviser les attentes avec de nouveaux objectifs

Un hiver de l'IA n'est pas nécessairement une catastrophe. C'est même tout le contraire : on peut le considérer comme une opportunité de prendre du recul et de réfléchir aux différents problèmes qui se sont posés durant la course effrénée vers un résultat extraordinaire. Le premier hiver de l'IA aura été bénéfique dans deux grands domaines de réflexion (sans compter des bénéfices mineurs ailleurs) :

» **La programmation logique :** cette discipline consiste à mettre une série de propositions sous une forme logique (exécutée en tant qu'application) qui exprime des faits et des règles concernant un domaine particulier. Les langages de programmation qui se fondent sur ce paradigme sont par exemple Prolog, Answer Set Programming (ASP) et Datalog. Cette forme de programmation par règles est la technologie sur laquelle reposent les systèmes experts.

» **Le raisonnement de sens commun** : ce domaine de pensée utilise une méthode simulant la capacité humaine de prédire le résultat d'une séquence d'événements d'après les propriétés, la finalité, les

intentions et le comportement d'un objet particulier. Le raisonnement de sens commun est une composante essentielle en IA car il exerce un impact sur diverses disciplines : vision par ordinateur, manipulation robotique, raisonnement taxonomique, action et changement, raisonnement temporel et raisonnement qualitatif.

Le second hiver de l'IA a apporté d'autres changements qui ont servi à susciter l'attention qui lui est accordée aujourd'hui. Ces changements ont été les suivants :

» **L'utilisation d'un matériel informatique courant :** à un moment donné, les systèmes experts et autres applications de l'IA fonctionnaient sur du matériel spécial. En effet, les matériels courants n'offraient pas la puissance de traitement ni la quantité de mémoire nécessaires. Or, ces systèmes sur mesure se sont révélés coûteux à entretenir, difficiles à programmer, et extrêmement précaires lorsque la situation était inhabituelle. Les matériels courants, conçus pour une utilisation généraliste, sont moins sensibles aux problèmes qui se posent lorsque l'on dispose d'une solution pour laquelle il s'agit de trouver un problème (voir la section « Créer des solutions pour lesquelles on cherchera un problème », plus loin dans ce chapitre).

» **La perception d'un besoin d'apprentissage** : les systèmes experts et autres formes initiales d'IA devaient être programmés spécialement pour chaque besoin à satisfaire, ce qui les rendait extrêmement peu flexibles. Il est devenu évident que les ordinateurs allaient devoir devenir capables d'apprendre de l'environnement, des capteurs et des données fournies.

» **La création d'un environnement flexible :** les systèmes qui accomplissaient des tâches utiles entre

le premier et le second hivers de l'IA fonctionnaient de façon rigide. Quand les données entrantes ne correspondaient pas exactement aux attentes, ces systèmes pouvaient produire des résultats contenant des erreurs grotesques. À l'évidence, il fallait que les prochains systèmes sachent s'adapter aux données réelles, qui comportent beaucoup d'erreurs, sont incomplètes et sont souvent formatées de manière incorrecte.

» **Le recours à de nouvelles stratégies :** imaginons que vous travailliez pour le gouvernement et que vous ayez promis toutes sortes de choses remarquables reposant sur l'IA, mais qu'aucune ne semble pouvoir se matérialiser. C'est ce qui s'est produit avec le second hiver de l'IA : différents gouvernements ont essayé par divers moyens de faire des promesses de l'IA une réalité. Quand les stratégies courantes ne donnaient visiblement pas satisfaction, ces mêmes gouvernements cherchaient d'autres moyens de faire progresser l'informatique, ce qui a parfois permis de produire des résultats intéressants, notamment dans le domaine de la robotique.

L'idée est que les hivers de l'IA ne sont pas nécessairement une mauvaise chose pour l'IA. Ce sont d'importantes occasions de prendre du recul et de faire le point sur le progrès (ou l'absence de progrès) des stratégies courantes. Certes, il est difficile de valoriser au mieux ces moments de réflexion quand on court désespérément vers le prochain succès en vue.



À propos des hivers de l'IA et du renouveau de l'IA qui leur succède, avec des idées et des objectifs révisés, il convient de songer à cet adage trouvé par le chercheur et futurologue Roy Charles Amara (qu'on appelle aussi la loi d'Amara) : « Nous avons tendance à surestimer l'impact

d’une technologie à court terme et à sous-estimer cet impact à long terme. » Après tout le battage et la désillusion, il arrive toujours un moment où les gens sont incapables de percevoir clairement l’impact à long terme d’une nouvelle technologie et de comprendre les révolutions qu’elle va engendrer. En tant que technologie, l’IA va perdurer et elle va changer notre monde, pour le meilleur ou pour le pire, quel que soit le nombre d’hivers qu’il lui faudra encore affronter.

Créer des solutions pour lesquelles on cherchera un problème

Deux personnes contemplent un amas de câblages, de roues, de morceaux de métal et d’éléments bizarres. La première personne demande : « Qu’est-ce que ça fait ? » et la seconde lui répond : « Qu’est-ce que ça ne fait pas ? » Or, l’invention qui semble pouvoir tout faire se révèle ne rien faire du tout. On trouve dans les médias de nombreux exemples de solutions en quête d’un problème. De façon risible, tout le monde a déjà trouvé une solution dont on cherchait quel problème elle était censée résoudre. Ces solutions-là n’ont pas d’avenir, même quand elles sont efficaces, parce qu’elles ne répondent pas à un besoin pressant. Les sections qui suivent étudient plus en détail la solution d’IA en quête de problème.

Définir un gadget

Dès qu’il s’agit d’IA, le monde est plein de gadgets. Certains de ces gadgets sont réellement utiles, mais beaucoup ne le sont pas et quelques-uns se situent entre

ces deux extrêmes. Alexa, par exemple, présente un certain nombre de fonctionnalités, mais également tout un ensemble de gadgets qui vous laissent perplexe quand vous essayez de les utiliser. Cet article de John Dvorak peut sembler un peu trop pessimiste, mais il donne à réfléchir sur ce sujet : https://www.pcmag.com/commentary/354629/just-say-no-to-amazons-echo-show.

Un gadget de l'IA est une application qui, à première vue, semble faire quelque chose d'intéressant, mais qui se révèle finalement incapable d'exécuter des tâches utiles. Les principaux critères pour savoir si une application de l'IA est autre chose qu'un gadget sont les suivants (leurs initiales forment l'acronyme CREEP, qui désigne en anglais ce qui est louche, et doit donc être évité) :

> » **Rentable (*Cost effective*) :** pour qu'une application de l'IA trouve un débouché, il ne faut pas qu'elle se révèle plus coûteuse que les solutions existantes. Les gens cherchent à faire des affaires. Personne n'est prêt à payer plus pour un avantage similaire.
>
> » **Reproductible** : les résultats d'une application doivent être reproductibles, même si les circonstances dans lesquelles la tâche doit être exécutée ne sont plus les mêmes. Contrairement aux solutions procédurales, l'IA est censée s'adapter, apprendre par la pratique, ce qui signifie que l'on place plus haut la barre.
>
> » **Efficiente** : dès qu'une solution consomme beaucoup de ressources, les utilisateurs vont voir ailleurs. Les entreprises, en particulier, cherchent plus que jamais à consommer le moins de ressources possible dans chaque activité.
>
> » **Efficace :** offrir simplement une solution pratique, rentable et efficiente ne suffit pas : il faut encore que cette solution réponde pleinement à un besoin. Une

solution efficace permet l'automatisation d'une tâche sans qu'il faille constamment revérifier les résultats ou améliorer le processus.

» **Pratique :** une application utile doit apporter un avantage pratique. Elle doit répondre à un besoin de l'utilisateur, comme par exemple la détermination de l'itinéraire à suivre sur une carte, ou le rappel en temps et en heure des médicaments à prendre.

Éviter l'infopublicité

Quand les utilisateurs potentiels de votre application la trouvent surprenante, vous pouvez être sûr que vous allez droit à l'échec. Assez curieusement, les applications qui ont les meilleures chances de succès sont celles dont l'objet est évident dès le départ. C'est le cas, par exemple, d'une application de reconnaissance vocale : vous parlez à la machine, et elle exécute une tâche utile. Vous n'avez pas besoin de chercher à convaincre les gens de l'utilité d'un tel logiciel. Vous trouverez dans ce livre un certain nombre d'applications vraiment utiles. Aucune ne nécessite un démarchage effréné. Si les gens commencent à vous demander à quoi sert votre produit, c'est qu'il est temps de repenser votre projet.

Savoir dans quels cas l'humain fait mieux

Ce chapitre fait le point sur la bonne utilisation de l'IA, sur ce que l'IA peut gérer ou non, et sur ce que nous faisons mieux que l'IA. Tout ce qui nécessite de l'imagination, de la créativité, la capacité de discerner où se trouve la vérité, la prise en compte d'une opinion ou le lancement d'une idée, tout cela est plutôt l'affaire des humains. De façon

assez curieuse, les limites de l’IA laissent beaucoup de champ libre à l’être humain, notamment dans un certain nombre de domaines auxquels nous ne pouvons pas encore accéder parce que nous sommes trop pris par des tâches répétitives et ennuyeuses qui pourraient facilement être confiées à l’IA.

Dans l’avenir, les meilleures applications de l’IA, et celles qui connaîtront le plus large succès, seront celles destinées à assister les gens, plutôt qu’à les remplacer. Il est exact que les robots remplaceront les humains là où des tâches doivent être exécutées dans des conditions difficiles, mais il reviendra aux humains de prendre les décisions qui permettront d’éviter que ces conditions deviennent plus difficiles encore, ce qui signifie qu’un humain, installé dans un endroit sécurisé, pilotera le robot. Ce sera une collaboration entre la technologie et les humains.

QUAND ON REPENSE À LA RÉVOLUTION INDUSTRIELLE

La collaboration entre l'IA et l'être humain ne se matérialisera pas du jour au lendemain. Par ailleurs, les nouvelles activités auxquelles les humains pourront se consacrer ne vont pas se généraliser de façon immédiate. Cependant, il n'est pas concevable que les gens restent inactifs, à attendre d'être servis par une machine. Nous continuerons à remplir des tâches variées. Naturellement, l'idée que les machines allaient remplacer les hommes a couru chaque fois qu'un grand bouleversement s'est produit dans l'histoire, et la révolution industrielle a été le plus récent et le plus violent de ces bouleversements (voir http://www.historydoctor.net/Advanced%20Placement%20World%20History/40.%20The_Industrial_revolution.htm). Il est des choses que les humains feront toujours mieux qu'une IA, et nous pouvons avoir la certitude que nous aurons toujours notre place dans la société. Il faut seulement espérer que le bouleversement à venir sera moins violent que l'a été la révolution industrielle.

Chercher la solution simple

Lorsqu'il s'agit de développer des applications de l'IA, le mieux est de garder à l'esprit le principe KISS (*Keep It Simple, Stupid*). Pour plus de détails sur le principe KISS, consultez la page https://www.techopedia.com/definition/20262/keep-it-simple-stupid-principlekiss-principle, mais il faut surtout retenir que toute solution doit être la plus simple possible. Ce principe a déjà été maintes fois formulé sous différentes formes. Le rasoir

d'Occam est sans doute l'exemple le plus fameux (https://science.howstuffworks.com/innovation/scientific-experiments/occams-razor.htm).

Naturellement, la question se pose de savoir pourquoi le principe KISS est aussi important. La réponse la plus immédiate est que la complexité entraîne l'échec : plus les éléments constituant un système sont nombreux, plus ce système risque d'avoir une défaillance. Ce principe trouve son fondement dans les mathématiques, et il est facilement démontrable.



Dans le domaine des applications, d'autres principes entrent en jeu. Une application est généralement considérée comme un moyen en vue d'une fin. Ce qui intéresse les gens, c'est la fin. On ne se soucie pas réellement de l'application. Si l'application pouvait devenir invisible, cela conviendrait parfaitement à l'utilisateur, qui ne se soucierait que du résultat. Les applications simples sont les plus faciles à utiliser, elles ont tendance à ne plus être visibles et elles ne nécessitent pas des instructions complexes. Les meilleures applications sont les plus évidentes. Si votre solution d'IA doit reposer sur toutes sortes d'interactions complexes, demandez-vous si vous ne devriez pas reprendre votre étude à zéro pour trouver quelque chose de mieux.

Chapitre 16
Voir l'IA dans l'espace

DANS CE CHAPITRE

» **Explorer l'univers**

» **Commencer l'exploitation minière dans l'espace**

» **Chercher de nouveaux endroits à explorer**

» **Développer des structures dans l'espace**

Depuis la nuit des temps, les hommes observent le ciel. Les noms des constellations et des étoiles viennent souvent de la mythologie grecque ou d'autres cultures de l'Antiquité (selon l'endroit où vous habitez). La Grande Ourse elle-même porte différents noms, et c'est en la groupant avec d'autres étoiles qu'on peut imaginer la forme d'un ours (voir http://newsok.com/article/3035192). Beaucoup de gens s'intéressent aux étoiles et adorent les observer, ce qui explique qu'un peu partout dans le monde, on ait cherché à identifier des formes dans le ciel. Quand l'homme est devenu capable de voyager dans l'espace, c'est tout l'univers qui a pris une nouvelle dimension, dont il sera question dans ce chapitre. En effet, l'IA permet de mieux percevoir l'univers et de le percevoir autrement.

L'homme a fini par expérimenter la vie dans l'espace (comme sur la Station spatiale internationale, https://www.nasa.gov/mission_pages/station/main/index.html) et par visiter d'autres lieux, comme

la Lune. L'homme a aussi commencé à travailler dans l'espace. Bien sûr, diverses expérimentations ont abouti à des productions qui n'étaient pas possibles ailleurs. Une entreprise s'est même spécialisée dans cette activité, il s'agit de Made In Space (http://madeinspace.us/). En dehors de ces activités, l'utilisation de robots et d'IA spécialisée permet l'exploitation minière de toutes sortes de matériaux dans l'espace. En 2015, le Congrès américain a promulgué une législation qui rend cette activité financièrement viable en garantissant aux entreprises le droit de vendre ce qu'elles extraient. Ce chapitre examine aussi le rôle de l'IA dans l'activité minière spatiale.

L'univers recèle une infinité de secrets. Une découverte récente est l'existence des exoplanètes, des planètes extérieures au système solaire (voir https://www.nasa.gov/feature/jpl/20-intriguing-exoplanets). L'existence d'exoplanètes signifie que les humains pourraient peut-être finir par découvrir une vie ailleurs, mais même pour trouver les exoplanètes, il fallait l'IA. La manière dont l'IA peut rendre visibles ces possibilités est véritablement admirable.

Vivre et travailler dans l'espace est une chose, mais y passer ses vacances en est clairement une autre. Dès 2011, la possibilité a été étudiée de construire un hôtel en orbite terrestre basse (http://mashable.com/2011/08/17/commercial-spacestation/) ou sur la Lune. En orbite terrestre basse, cela semble aujourd'hui envisageable, mais l'hôtel sur la Lune attendra encore. Toujours est-il que l'IA permettra un jour d'habiter, de travailler et même, de partir en vacances dans l'espace grâce aux structures spécialisées qui sont présentées dans ce chapitre.

Observer l'univers

C’est à un fabricant de lunettes hollandais, Hans Lippershey, que l’on attribue l’invention du télescope au tout début du XVIIe siècle. Peu de temps auparavant, il y avait les *lunettes d’approche*. En réalité, la paternité du télescope est controversée (voir https://www.space.com/21950-who-invented-the-telescope.html). Des scientifiques comme l’astronome italien Galilée scrutaient déjà le ciel à l’aide d’un instrument d’optique. Le principe du télescope est donc plus ancien. Au cours du temps, les appareils d’observation de l’espace sont devenus plus volumineux, plus complexes, et maintenant certains sont eux-mêmes dans l’espace.



La raison pour laquelle on cherche à installer des télescopes dans l’espace est que l’atmosphère terrestre rend impossible l’obtention d’images nettes de ce qui est très éloigné. Le télescope Hubble est un des premiers et des plus fameux télescopes spatiaux. Comme on le verra dans les sections suivantes, l’utilisation des télescopes modernes fait appel à l’IA à plus d’un titre, par exemple pour programmer le temps d’utilisation du Hubble (voir http://ieeexplore.ieee.org/document/63800/?reload=true).

Voir clair pour la première fois

Un moyen d’éviter le problème de l’atmosphère terrestre consiste à installer le télescope dans l’espace. Cependant, cette approche est coûteuse, et la maintenance peut devenir un cauchemar. Les gens qui observent le ciel ont généralement besoin d’une autre solution, par exemple un télescope équipé d’un système à laser capable de déformer le miroir de manière à pallier le flou résultant de l’atmosphère (voir https://www.space.com/8884-telescope-laser-vision-heavens-blurry.html).



Imaginez que vous deviez calculer l'effet de flou exercé par l'atmosphère terrestre plusieurs milliers de fois par seconde, en vous référant à la lumière d'un laser. Le seul moyen d'effectuer un nombre de calculs aussi considérable pour ensuite déplacer les actionneurs sur le miroir de façon précise consiste à recourir à l'IA. Un exemple parmi d'autres d'utilisation de l'IA dans le domaine de l'optique adaptative est présenté dans cet article : https://www.spiedigitallibrary.org/conference-proceedings-of-spie/2201/1/Artificial-intelligence-system-and-optimized-modal-control-for-the-ADONIS/10.1117/12.176120.short ? SSO=1. Pour découvrir la manière dont les réseaux neuronaux sont utilisés dans les systèmes d'optique adaptative.

Pour obtenir des optiques meilleures encore, les télescopes du futur utiliseront une correction 3D des effets de flou grâce à des optiques adaptatives multiconjuguées (http://eso-ao.indmath.uni-linz.ac.at/index.php/systems/multi-conjugate-adaptive-optics.html). Cette nouvelle technologie corrigera le champ de vision étroit qu'offrent les télescopes actuels, mais elle nécessitera un contrôle plus grand encore (et plus précis) des multiples niveaux d'actionneurs agissant sur plusieurs miroirs. Les nouveaux télescopes, notamment le télescope géant Magellan, le Télescope de Trente Mètres et le Télescope géant européen, seront basés sur cette technologie, ce qui justifiera l'investissement réalisé (supérieur à un milliard de dollars).

Trouver de nouveaux endroits où aller

Jusqu'au XVIIIe siècle, les hommes étaient rivés à la surface de la Terre mais rêvaient en contemplant le ciel. Certains tentaient toutes sortes d'expériences douteuses, comme sauter du sommet d'une tour (voir https://jkconnectors.com/news/the-history-of-aviation-part-1/), mais avant l'invention des ballons à air chaud, toute forme de vol dans les airs semblait inaccessible. Cependant, les recherches se sont poursuivies et aujourd'hui, les hommes cherchent encore, et ils cherchent aussi de nouveaux endroits où aller.



L'idée de chercher de nouveaux lieux à explorer ne s'est vraiment matérialisée que le 20 juillet 1969 avec la première expédition sur la Lune (voir https://www.nasa.gov/mission_pages/apollo/apollo11.html). Auparavant, on pouvait chercher, certes, mais on ne pouvait rien atteindre. Depuis, l'homme cherche toutes sortes d'endroits où se rendre et il en a atteint quelques-uns, comme Mars et la comète Rosetta. Chacune de ces explorations sert à stimuler le désir d'aller découvrir d'autres endroits de l'espace. Plus important, aucune n'aurait pu avoir lieu sans les calculs mathématiques complexes qu'une IA peut exécuter.

Pour ce genre de découvertes, il fallait des télescopes, mais de plus en plus, la NASA et d'autres organismes adoptent des méthodes différentes, comme l'IA. En l'occurrence, l'apprentissage machine a permis de localiser une huitième planète autour de Kepler 90. Naturellement, quand on découvre tous ces endroits à explorer, le problème est de déterminer si nous pourrons vraiment les atteindre. Voyager 1, la sonde la plus éloignée de la Terre, n'a atteint l'espace interstellaire que récemment (https://www.space.com/26462-voyager-1-interstellar-space-confirmed.html). Ses moteurs sont endommagés mais restent utilisables. Cependant, à 20 milliards de kilomètres, Voyager n'est qu'à 0,0022 année-lumière d'ici,

et il lui a fallu 40 ans pour y parvenir. Kepler 90 étant à 2545 années-lumière, il semble impossible de l'atteindre sans disposer d'une technologie radicalement nouvelle, qui sera peut-être créée un jour avec l'aide de l'IA.



Heureusement, il y a dans notre système solaire toutes sortes d'endroits que l'on pourrait atteindre. L'*Encyclopædia Britannica*, par exemple, recommande de visiter des endroits comme le bassin Caloris sur Mercure (voir https://www.britannica.com/list/10-places-to-visit-in-the-solar-system). Nous vous invitons aussi à consulter TravelTips4Life (http://www.traveltips4life.com/15-places-we-want-tovisit-in-outer-space/), qui recommande la Station spatiale internationale en guise de première étape.

Étudier l'évolution de l'univers

Les humains contemplent l'univers depuis longtemps, mais n'ont encore aucune idée réelle de ce qu'il est, ils savent simplement qu'ils y vivent. Bien sûr, les observations continuent, mais la nature même de l'univers reste une grande inconnue. Récemment, des scientifiques ont commencé à utiliser l'IA pour déterminer soigneusement le mouvement de différentes parties de l'univers en vue de découvrir la manière dont il fonctionne (voir https://www.sciencedaily.com/releases/2012/09/120924080307.htm). Le modèle Lambda Cold Dark Matter (LCDM) nous permettra de le comprendre un peu mieux. Cependant, nous sommes encore bien loin de pouvoir commencer à trouver les réponses à nos nombreuses questions.

Créer de nouveaux principes scientifiques

Au bout du compte, il faut que les recherches effectuées pour en savoir plus sur l'espace, le système solaire, la galaxie et l'univers soient payantes. Autrement, personne ne continuera à les financer. Les hivers de l'IA évoqués au Chapitre 15 sont un exemple de ce qui arrive à une technologie, aussi prometteuse qu'elle puisse être, quand elle ne répond pas aux attentes. Par conséquent, compte tenu de la longue histoire de l'exploration de l'espace, il faut bien que l'humanité en tire quelque profit. Dans la plupart des cas, ce profit prend la forme de principes scientifiques : un progrès dans la compréhension du fonctionnement des choses. En appliquant les leçons tirées de l'exploration de l'espace et des voyages dans le cosmos, on peut vivre mieux sur terre. En outre, il arrive souvent que des technologies utilisées dans l'espace débouchent sur des produits de consommation courante.

Prenons l'exemple d'une exploration : l'alunissage d'Apollo 11. Nous ressentons encore les effets de la révolution des technologies qui s'est produite durant la préparation de cette mission. Par exemple, le besoin d'économiser de l'espace a conduit les autorités américaines à consacrer de lourdes dépenses à des technologies comme les circuits intégrés (CI) que nous utilisons aujourd'hui sans même y penser (voir https://www.computerworld.com/article/2525898/app-development/nasa-s-apollo-technology-has-changed-history.html). Selon les sources, chaque dollar investi par le gouvernement américain dans les travaux de recherche de la NASA rapporte aujourd'hui aux Américains 7 à 8 dollars de biens et de services.

Cependant, au-delà de la mise au point des capsules actuelles et de leurs composants, la course à l'espace a débouché sur une nouvelle technologie. Le film *Les Figures de l'ombre*, par exemple (https://www.amazon.com/exec/obidos/ASIN/B01LTI1RHQ/datacservip0f-20/), présente une vision de la NASA différente de celle que peuvent avoir la plupart des gens : pour tous ces calculs, il faut beaucoup de puissance informatique. Dans le film, on voit l'évolution des calculs de la NASA, des calculateurs humains aux calculateurs électroniques. Cependant, si vous êtes attentif, vous vous rendrez compte que l'ordinateur travaille finalement en collaboration avec l'être humain, de manière similaire à celle dont l'IA travaillera avec les humains à mesure que notre connaissance de l'univers progressera.



Aujourd'hui, nos données sur l'espace proviennent de partout. Elles nous permettent d'élaborer de nouveaux principes scientifiques concernant des choses que nous ne voyons même pas, comme les *espaces noirs* (des régions de l'espace dotées d'une masse mais sans présence visible) et l'énergie noire (une forme d'énergie inconnue et non identifiée qui s'oppose aux effets de la gravitation entre les corps dans l'espace). L'étude de ces entités invisibles nous permet d'acquérir de nouvelles connaissances sur la manière dont agissent les forces sur notre propre planète. Cependant, les chercheurs sont tellement submergés de données qu'ils doivent recourir à l'IA rien que pour en comprendre une petite partie (voir https://www.theverge.com/2017/11/15/16654352/ai-astronomy-spaceexploration-data). L'avenir de l'espace et la manière dont nous utilisons les technologies créées pour l'espace dépendent de l'usage que nous pouvons faire de toutes ces données que nous recueillons, et c'est là que l'IA est indispensable.

Extraire du minerai dans l’espace

Les activités minières dans l’espace sont un sujet dont se sont copieusement abreuvés les médias ainsi que les scientifiques. Des films comme *Alien* (https://www.amazon.com/exec/obidos/ASIN/B001AQO3QA/datacservip0f-20/) donnent une petite idée de ce à quoi pourrait ressembler un futur vaisseau destiné à ces activités (en espérant qu’il n’y aura pas d’extraterrestres hostiles dans la réalité). En fait, des entreprises comme Deep Space Mining étudient déjà les conditions dans lesquelles des activités minières spatiales seraient possibles. Ce qui est surprenant, c’est que ces entreprises minières s’intéressent à la possibilité de trouver de l’eau, un élément très courant sur terre mais relativement difficile à trouver dans l’espace. Les sections qui suivent donnent un aperçu plus complet des aspects intéressants des activités minières spatiales.

NE PAS NÉGLIGER LA CRITIQUE

Rares sont ceux qui savent apprécier le rôle des critiques dans la société : ces personnes qui excellent pour trouver le mauvais côté des choses, qui soulignent le revers de la médaille ou qui repèrent des nids-de-poule sur toutes les routes, ces rouspéteurs et ces grincheux qui sont si souvent stigmatisés dans les médias. Or, la critique a un important rôle à jouer dans le domaine de l'IA appliquée aux activités spatiales. Orientée comme il se doit, elle débouche sur la planification à long terme qui semble faire défaut chez les membres les plus optimistes d'une équipe. Alors que tous les autres dirigent leur attention sur des solutions créatives aux problèmes existants, la critique se soucie des problèmes futurs, qui ont effectivement une grande importance pour les applications de l'IA, entre autres pour celles utilisées dans les activités minières.

Une IA appliquée à des activités spatiales devra davantage être indépendante que si elle s'appliquait à des activités terrestres. Concernant les différentes sondes déjà lancées dans l'espace, il devient évident que planifier ce qui n'est pas planifiable est maintenant une chose obligatoire, et non pas une sympathique fonctionnalité optionnelle. Dans l'espace, une IA devra apprendre de l'environnement dans lequel elle se trouve et définir des solutions à des problèmes auxquels les développeurs humains n'auront peut-être pas pensé, par exemple des effets gravitationnels inattendus, des défaillances matérielles imprévues, une pénurie de pièces de rechange, etc. Le matériel aussi doit être confronté à des défis qu'une IA destinée aux activités spatiales n'est pas encore en mesure d'affronter, comme des tentatives, par des pirates, de voler les cargaisons. Un esprit tourné vers la critique pourra apporter une précieuse contribution dans ce domaine et sera donc un élément essentiel dans chaque équipe.

L'encadré « Comprendre l'orientation de l'apprentissage » dans le Chapitre 13 recèle également d'importantes leçons

concernant l'IA appliquée aux activités spatiales. Une de ces leçons concerne la futilité, en d'autres termes, la conscience que dans certaines situations, il n'y a pas de gagnant. Une IA spatiale pourrait donc prendre des contre-mesures pour éviter les dommages, plutôt que nécessairement tenter de résoudre un problème insoluble. Dans l'espace, il y aura une infinité d'inconnues, ce qui signifie que l'intervention humaine sera nécessaire, mais cette intervention demandera peut-être plusieurs mois. Dans cette attente, l'IA spatiale devra savoir comment préserver le potentiel de capacités opérationnelles. Les considérations que l'on peut lire sur la page https://worldbuilding.stackexchange.com/questions/66698/what-issueswould-an-aiasteroid-mining-stations-have-to-be-pre-pared-for ne constituent qu'un petit échantillon de la formidable quantité de problèmes auxquels une IA spatiale pourrait être confrontée.

Récolter de l'eau

L'eau recouvre environ 71 % de notre planète. Il y a même tellement d'eau qu'il nous est souvent difficile d'éviter qu'elle se répande là où nous ne voulons pas qu'elle soit. La Terre fait cependant figure d'exception dans ce domaine. Il n'y a pas surabondance d'eau dans l'espace. Naturellement, vous vous demandez peut-être pourquoi on pourrait avoir besoin de trouver de l'eau dans l'espace, si ce n'est pour les besoins des astronautes et éventuellement des plantes. Il se trouve que les deux éléments dont l'eau (H2O) est constituée, l'hydrogène et l'oxygène, sont aujourd'hui deux composants essentiels du carburant des fusées (voir https://www.nasa.gov/topics/technology/hydrogen/hydrogen_fuel_of_choice.html). Par conséquent, cette grosse boule de glace sale dans le ciel

pourrait bien devenir un jour une station de ravitaillement en carburant.

UTILISER DES DRONES ET DES ROBOTS DANS L'EXPLOITATION MINIÈRE

On ne peut pas déterminer ce que contient un astéroïde tant qu'on ne s'en est pas considérablement approché. En outre, le nombre d'astéroïdes qu'il faut explorer avant de pouvoir trouver quelque chose d'exploitable est significativement élevé : bien plus élevé que le nombre d'astéroïdes que des pilotes humains pourraient jamais explorer. Par ailleurs, s'approcher d'un objet qui tourne sur lui-même d'une façon particulière et qui possède des caractéristiques étranges n'est pas sans danger. Pour toutes ces raisons, les explorations d'astéroïdes aux fins d'exploitation minière seront généralement effectuées par diverses sortes de drones autonomes, qui iront d'un astéroïde à un autre à la recherche des matériaux voulus. Quand un drone trouvera un minerai, il alertera une station centrale en lui transmettant les informations nécessaires, à savoir, la localisation précise de l'astéroïde et ses caractéristiques.

Un robot sera alors envoyé vers l'astéroïde. On s'imagine généralement que l'exploitation minière se fera sur place, mais ce serait à la fois coûteux et dangereux. Une autre possibilité serait de déplacer d'abord l'astéroïde vers un lieu plus sûr, par exemple en orbite autour de la Lune. Il s'agit même là d'un projet réel pour lequel un financement existe (voir https://www.outerplaces.com/science/item/1332-nasa-tore-ceive-100m-budget-for-asteroid-capture). Le déplacement de l'astéroïde sera assuré par des robots, et l'exploitation minière sera probablement effectuée par d'autres robots. Les humains interviendront peut-être pour la réparation des robots, et sans doute pour la surveillance des activités des robots et des drones. Ce seront des activités minières moins dangereuses, moins polluantes et plus intéressantes que celles que l'on pourrait mener sur Terre.

Obtenir des terres rares et autres métaux rares

Les activités minières ont toujours été salissantes, mais certaines le sont plus que d'autres et les terres rares relèvent de cette catégorie. L'exploitation des terres rares est si polluante (voir http://earthjournalism.net/stories/the-dark-side-of-renewable-energy) qu'aux États-Unis, toutes les mines de terres rares ont été fermées jusqu'à ce que le gouvernement trouve nécessaire de rouvrir la mine de Mountain Pass, qui constitue une réserve stratégique, sachant que la Chine a pris le contrôle du marché des terres rares. Un des pires inconvénients de l'exploitation des terres rares est qu'elle se traduit par la prolifération de thorium radioactif dans les zones environnantes.

En raison du coût extrêmement élevé, à la fois en termes d'environnement et de main-d'œuvre, de l'extraction des terres rares aux États-Unis, la poursuite de l'exploitation de la mine de Mountain Pass est mise en doute. En fait, les Américains se disputent avec les Chinois pour les empêcher d'acheter la seule mine américaine.

Votre téléphone portable, votre iPad, votre voiture, votre téléviseur, aussi bien que les panneaux solaires et les éoliennes, fonctionnent grâce à ces matériaux très dangereux que sont les terres rares (pour quelques exemples d'utilisation, voir http://www.rareearthtechalliance.com/Applications/Electronics.html). La plupart des gens ne savent même pas que ces matériaux ne sont pas durables, compte tenu de la manière dont ils sont actuellement utilisés (http://www.pbs.org/wgbh/nova/next/physics/rare-earth-elements-in-cell-phones/). Les

problèmes que pose l'exploitation de ces minéraux constituent la meilleure raison de les extraire ailleurs que sur notre planète, là où les rejets toxiques ne seront plus un danger pour nous. En réalité, leur extraction ne devrait être que la première étape : c'est toute cette industrie qui devrait être délocalisée loin de la Terre (en effet, elle est extrêmement polluante).



L'IA est essentielle dans les efforts pour trouver de meilleures sources de terres rares qui ne pollueront pas notre planète. Une curiosité remarquable que présentent les terres rares est le fait que la Lune soit susceptible de les fournir en quantités significatives (voir https://www.washingtonpost.com/national/health-science/moon-draws-growing-interest-as-a-potentialsource-of-rare-minerals/2012/01/30/gIQAqHvUuQ_story.html ?utm_term=.828c9cb19a34). En réalité, un certain nombre de dirigeants politiques voient maintenant l'extraction des terres rares sur la Lune comme un besoin stratégique (voir https://sservi.nasa.gov/articles/is-miningrare-minerals-on-the-moon-vital-to-national-security/). Le problème est que les efforts pour découvrir précisément de quoi la Lune est faite n'ont pas été entièrement couronnés de succès jusqu'à présent, or il est important de savoir à quoi s'attendre. Le Moon Minerology Mapper (https://www.jpl.nasa.gov/missions/moon-mineralogy-mapper-m3/) n'est qu'un exemple des efforts pour découvrir la composition de la Lune. Par ailleurs, pour pouvoir exploiter les terres rares et en faire des produits utiles, il faudrait disposer d'une source d'eau, et il semblerait qu'il y ait de l'eau sur la Lune. Pour les sondes, les robots, l'analyse de données et toute la planification requise, l'IA est indispensable, car les

problèmes sont considérablement plus compliqués que ce que l'on pourrait penser.

Trouver de nouveaux éléments

Le tableau périodique des éléments a fait l'objet d'un certain nombre de mises à jour au cours du temps. Quatre nouveaux éléments y ont même fait leur apparition en 2016 (voir https://www.sciencenews.org/blog/science-ticker/four-newestelements-periodic-table-get-names). Cependant, pour trouver ces quatre éléments, il a fallu au minimum une centaine de scientifiques et le recours à une IA avancée, sachant que ces éléments ne survivent généralement qu'une fraction de seconde en laboratoire. Curieusement, l'espace pourrait constituer un environnement dans lequel ces nouveaux éléments existeraient naturellement, plutôt que de disparaître en une fraction de seconde parce que les protons, dans le noyau, se repoussent mutuellement.



Comme on le voit ici, on trouve encore de nouveaux éléments à ajouter au tableau périodique, et de façon quasi certaine, l'espace nous en fournira d'autres. Les supernovae et autres phénomènes spatiaux peuvent permettre la réplication d'éléments que les scientifiques créent en utilisant des accélérateurs de particules ou des réacteurs. En fait, les physiciens des particules travaillent avec l'IA depuis les années 1980. Vous serez peut-être surpris d'apprendre qu'on a déjà trouvé un élément qui n'existe que dans l'espace, le technétium.

La combinaison d'éléments donne de nouveaux matériaux. L'IA est aussi ce qui permet aux chimistes de trouver de nouvelles manières de combiner des éléments pour obtenir de nouveaux cristaux intéressants (voir https://www.sciencedaily.com/releases/2016/

09/160921084705.htm). Une fois, les chercheurs ont découvert 2 millions de nouvelles sortes de cristaux en utilisant seulement quatre éléments, mais ces découvertes ont nécessité le recours à l'IA. Imaginons ce qui se produira dans l'avenir quand les scientifiques commenceront à ouvrir la porte à l'IA et à l'apprentissage profond (lesquels pourront déterminer si les cristaux obtenus sont réellement exploitables).

Améliorer la communication

Une entreprise spatiale aussi complexe que ce genre d'activité minière suppose le recours à des communications avancées. Même si les sondes et les robots utilisés pour l'exploitation minière utilisent des capacités d'apprentissage profond pour gérer la plupart des incidents mineurs et certains incidents majeurs qui se produiront au cours du processus, il reviendra toujours aux humains de résoudre des problèmes que l'IA ne pourra pas résoudre. Attendre des heures avant de découvrir l'existence d'un problème, puis passer encore des heures à essayer d'en déterminer la source, serait catastrophique pour l'exploitation minière dans l'espace. Les techniques actuelles de communication devront être modernisées, et aussi curieux que cela puisse paraître, l'IA interviendra là encore.



La radio cognitive (voir http://ieeexplore.ieee.org/document/5783948/) fait appel à l'IA pour prendre automatiquement des décisions concernant l'amélioration de l'efficience de la radio de diverses manières. L'opérateur humain n'a pas à s'inquiéter de savoir précisément comment le signal est transmis d'un endroit à un autre : il le fait simplement de la meilleure manière possible. Dans bien des cas, la radio cognitive atteint son objectif grâce à la partie inutilisée ou sous-utilisée du spectre, mais elle peut aussi l'atteindre

grâce à d’autres méthodes. En d’autres termes, les méthodes actuelles de contrôle des sondes comme celles présentées sur la page https://fr.wikipedia.org/wiki/Liste_des_missions_d%27exploration_spatiale_en_cours ne feront pas l’affaire dans l’avenir lorsqu’il faudra faire plus, en moins de temps et avec moins de spectre (en raison d’une charge accrue sur la communication).

Explorer de nouveaux endroits

L’espace est vaste. Il y a peu de chances que les humains puissent l’explorer en entier. Celui qui dit qu’il n’y a plus de frontières n’a jamais regardé le ciel. Même les auteurs de science-fiction semblent penser qu’il y aura toujours des endroits inexplorés dans l’univers. Bien sûr, si vous adhérez à la théorie des multivers (https://www.space.com/18811-multiple-universes-5-theories.html), le nombre d’endroits à explorer peut être infini. Le problème n’est même pas de trouver où aller : il est plutôt de savoir par où commencer. Les sections qui suivent vous éclairent sur le rôle de l’IA pour faire voyager des personnes vers les autres planètes, puis vers les étoiles.

Commencer par la sonde

Les hommes ont déjà entrepris d’envoyer des sondes partout, pour tout explorer. L’utilisation de sondes est même plus ancienne qu’on le croit souvent. Dès 1916, Robert H. Goddard, un pionnier américain dans le domaine des fusées, avait calculé qu’une fusée pourrait être envoyée sur la Lune au moyen d’une charge explosive visible depuis la Terre. Cependant, ce sont E. Burgess et C. A. Cross qui ont parlé de sonde pour la

première fois dans un article intitulé *The Martian Probe*, en 1952. Pour le commun des mortels, une sonde spatiale est un véhicule conçu pour quitter la Terre et aller explorer un autre lieu. La première sonde à avoir atterri sur la Lune était Luna 9, en 1966.

Aujourd'hui, les sondes ne se contentent pas d'atteindre une certaine destination. Une fois arrivées sur place, elles exécutent des tâches complexes, puis retransmettent les résultats par radio aux chercheurs qui sont sur Terre. Ainsi, la NASA a conçu la sonde Curiosity pour qu'elle permette de savoir si une vie microbienne a pu exister sur Mars. À cette fin, Curiosity est dotée d'un système informatique élaboré capable d'exécuter de lui-même un certain nombre de tâches. Dans bien des cas, il n'est pas question d'attendre l'arrivée des humains, et certains problèmes nécessitent une résolution immédiate. Curiosity produit une si grande quantité d'informations qu'elle peut avoir son propre blog, ses podcasts et son site Internet, comme on peut le voir sur la page https://www.nasa.gov/mission_pages/msl/index.html. Pour plus de détails sur la construction et les capacités de Curiosity, consultez la page https://www.space.com/17963-mars-curiosity.html.

Il n'est pas difficile d'imaginer quelle vaste quantité d'informations une sonde comme Curiosity peut transmettre. L'analyse des données de Curiosity, à elle seule, est comparable en volume de données à l'analyse des grandes données que réalisent des organisations comme Netflix ou Goldman Sachs. La différence est que le flux de données provient de Mars, et non pas des utilisateurs locaux, si bien que dans toute analyse de données il importe de prendre en compte le temps nécessaire à l'obtention de l'information. En fait, le délai de transmission entre la Terre et Mars est de 24 minutes. Sachant cela, on comprend que Curiosity et les autres

sondes doivent penser par elles-mêmes (https://www.popsci.com/artificial-intelligence-curiosityrover), même lorsqu'il s'agit d'effectuer certaines formes d'analyse.

Une fois que les données sont arrivées sur Terre, les scientifiques les stockent, puis les analysent. Même avec l'aide de l'IA, ce processus demande des années. À l'évidence, pour atteindre les étoiles, il faudra faire preuve de patience et il sera nécessaire de disposer d'une puissance de traitement plus importante encore que ce qui est envisageable actuellement. L'univers étant un espace désordonné, le recours aux sondes est essentiel, mais ces sondes peuvent avoir besoin d'être plus autonomes, ne serait-ce que pour pouvoir trouver les endroits adéquats à explorer.

Compter sur les missions robotiques

Pour en savoir davantage sur une planète, il y a peu de chances que nous puissions nous y rendre directement comme le font les personnages des romans et des films de science-fiction. Il est plus logique d'y envoyer des robots, qui permettront de déterminer s'il serait viable d'y envoyer des humains par la suite, sachant que l'utilisation des robots coûte moins cher et qu'on peut les faire voyager plus facilement. Des robots ont déjà été envoyés sur plusieurs planètes et sur plusieurs satellites au sein du système solaire, mais il semble que Mars soit leur destination favorite pour un certain nombre de raisons :

» Il est possible d'envoyer sur Mars une mission robotique tous les 26 mois.

» Mars se situe dans la zone habitable du système solaire, si bien qu'elle est susceptible d'être

colonisée.

» Un certain nombre de scientifiques pensent que la vie a existé sur Mars.

L'histoire d'amour entre l'humanité et Mars a commencé en octobre 1960, quand l'Union soviétique a lancé Marsnik 1 et Marsnik 2. Malheureusement, aucune de ces deux sondes n'a pu atteindre l'orbite terrestre, et encore moins la planète Mars. Ce sont ensuite les États-Unis qui ont essayé, avec les engins Mariner 3 en 1964 et Mariner 4 en 1965. Mariner 4 a pu nous transmettre 12 photos de la planète rouge. Depuis, d'innombrables sondes ont été envoyées sur Mars ainsi que toute une série de robots, et ces robots commencent à nous en révéler les secrets (toutefois, le taux de succès des expéditions sur Mars est inférieur à 50 %, selon cet article : https://www.space.com/16777-curiosity-rover-many-mars-missions.html). En dehors des sondes conçues pour effectuer des survols et observer Mars depuis l'espace, deux types de robots descendent sur Mars :

» **Le robot atterrisseur** : conçu pour se fixer quelque part et y exécuter des tâches relativement complexes.

» **Le rover** : robot mobile qui explore le terrain et permet ainsi d'étudier une plus grande surface.

ENVISAGER DES CIBLES POUR LA COLONISATION

D'après ce que l'on peut lire dans certaines revues, les scientifiques étudient d'ores et déjà des lieux que l'humanité serait susceptible de pouvoir coloniser dans l'avenir. Cette colonisation deviendra essentielle pour un certain nombre de raisons, dont la croissance démographique sur terre n'est pas la moindre. Naturellement, les activités industrielles et minières envisageables sur d'autres planètes entrent également en jeu. En outre, la possibilité de vivre sur une autre planète augmente nos chances de survie, au cas où un nouvel astéroïde tueur frapperait la Terre. Compte tenu de ces considérations, voici une liste des cibles communément envisagées pour la colonisation spatiale (la vôtre est peut-être différente) :

- la Lune
- Mars
- Europa
- Encelade
- Cérès
- Titan

Chacune de ces possibilités éventuelles s'accompagne d'exigences auxquelles une IA pourrait permettre de répondre. Pour coloniser la Lune, par exemple, il serait nécessaire d'utiliser des dômes. Une source d'eau serait également indispensable : il faudrait disposer d'eau en quantité suffisante pour pouvoir produire de l'oxygène, pour la respiration, et de l'hydrogène qui servirait de source de chaleur. Les sondes nous fourniront des informations, mais la modélisation de l'environnement nécessaire à cette colonisation demandera du temps et une capacité de traitement considérable, ici sur Terre, avant que l'humanité puisse commencer à déménager.

Vous trouverez une liste de robots atterrisseurs et de robots mobiles envoyés sur Mars depuis 1971 à l'adresse https://www.space.com/12404-mars-explored-landers-rovers-1971.html. La plupart de ces engins ont été envoyés par les États-Unis et l'Union soviétique, mais au moins un rover a été envoyé par les Anglais. Avec le progrès et la diffusion des techniques nécessaires pour réussir à faire atterrir des engins sur une autre planète, on peut s'attendre à voir bientôt d'autres pays participer à cette course vers Mars (même si ce n'est qu'à distance).



À mesure que les robots atterrisseurs et les robots mobiles se perfectionnent, l'IA devient de plus en plus nécessaire. La sonde Curiosity, par exemple, est dotée d'une IA relativement complexe qui lui permet de choisir, en toute autonomie, de nouvelles cibles à explorer, comme l'explique la page http://www.astronomy.com/news/2016/08/how-does-mars-rover-curiositys-new-ai-system-work. Cependant, il ne faudrait pas s'imaginer que cette IA remplace les chercheurs qui sont sur Terre. Ce sont toujours les chercheurs qui déterminent les propriétés des roches examinées par l'IA. Par ailleurs, il arrive qu'un chercheur passe outre l'IA et choisisse une cible différente. L'IA est là pour assister les scientifiques, et non pour les remplacer. Nous avons là un exemple de la manière dont l'homme et l'IA travailleront ensemble dans l'avenir.

Même si, à ce jour, toutes les expéditions robotiques réussies ont été financées par des fonds publics, les activités minières et commerciales dans l'espace nécessiteront un jour ou l'autre la participation du secteur privé. Google, par exemple, a lancé le concours Lunar XPRIZE pour la première entreprise commerciale sur la Lune, assorti d'un prix de 20 millions de dollars. Pour gagner, les candidats doivent réussir à faire alunir un engin robotique et à lui faire parcourir 500 mètres et retransmettre une vidéo à haute définition. Ce concours

est important, car le projet en question ne sera pas réalisé uniquement pour l'obtention du prix : ce sera le prélude d'autres entreprises.

Ajouter l'élément humain

L'homme est curieux de connaître d'autres lieux que la Terre. Bien sûr, pour le moment il n'est allé que sur la Lune, pour la première fois le 20 juillet 1969, dans le cadre de la mission Apollo 11. Depuis, l'homme a mis le pied sur la Lune à six reprises, la dernière fois le 7 décembre 1972 avec Apollo 17. La Chine, l'Inde et la Russie étudient également des projets d'expédition sur la Lune. Les Russes prévoient une expédition vers 2030. La NASA prévoit d'y retourner, mais rien n'est encore programmé dans ce sens.

Concernant la planète Mars, la NASA a lancé des projets. Cependant, l'homme ne devrait pas mettre le pied sur Mars avant les années 2030. Comme on peut l'imaginer, aussi bien la science des données que l'IA, l'apprentissage machine et l'apprentissage profond occuperont une place prépondérante dans toute entreprise visant à atteindre Mars. En raison de la distance et du contexte, d'importants concours seront nécessaires. En outre, revenir de Mars sera considérablement plus difficile que revenir de la Lune. Même le décollage sera plus difficile, compte tenu de la présence d'une atmosphère et d'une plus forte gravité sur Mars.



En 1968, Arthur C. Clark publiait son ouvrage *2001* : l'Odyssée de l'espace. Ce livre aura sans doute touché une corde sensible, sachant qu'il a inspiré un film et une série télévisée, sans compter trois autres livres. Dans cette histoire, l'ordinateur à programmation heuristique HAL 9000 (HAL, pour *Heuristically programmed ALgorithmic*) finit par devenir incontrôlable, par suite d'un

conflit dans les paramètres de sa mission. Il était censé, plus que toute autre chose, aider les voyageurs de l'espace à réussir leur mission, mais son objet implicite était aussi de leur éviter de devenir fous du fait de leur isolement. Il y a peu de chances qu'un ordinateur ressemblant à HAL soit utilisé dans les futurs vols spatiaux. Pour commencer, une IA programmée pour l'espace ne risque pas de tenir l'équipage dans l'ignorance des paramètres de la mission. Il ne fait aucun doute que les vols dans l'espace se feront avec le concours d'une IA, mais celle-ci sera conçue de façon plus pratique et plus triviale que le HAL 9000.

Construire des structures dans l'espace

À un moment donné, explorer l'espace ne suffira plus. La réalité des voyages spatiaux est que tout est situé si loin de tout le reste qu'il faudra des points de passage entre les destinations, et même avec ces points de passage, les expéditions exigeront de sérieux efforts. Toutefois, ces points de passage ont déjà une grande importance. Imaginons que l'exploitation minière sur la Lune commence dès demain. Il sera alors impératif de disposer d'un entrepôt en orbite terrestre basse, compte tenu du coût colossal que représentera le transfert d'un équipement minier et d'autres ressources depuis la surface de la Terre. Naturellement, un voyage en sens inverse sera également nécessaire pour rapporter sur terre les ressources extraites et les produits finis. Certains voudront aussi prendre leurs vacances dans l'espace, et les scientifiques utilisent déjà diverses structures pour poursuivre leurs investigations. Les sections qui suivent étudient les utilisations variées de ces structures aux fins

de permettre à l'humanité, depuis la planète Terre, de gagner les étoiles.

Passer ses premières vacances dans l'espace

Cela fait maintenant quelque temps que des entreprises promettent des vacances dans l'espace. Orbital Technologies a été parmi les premières à le faire en 2011, la date initialement prévue étant 2016 (voir http://www.smh.com.au/technology/sci-tech/spacevacation-orbiting-hotel-ready-for-guests-by-2016-20110818-1j0w6.html). L'idée était d'emmener sept personnes dans une fusée Soyouz (russe) pour un voyage de cinq jours. Ce n'est pas encore possible, mais cette vidéo montre quelle est la technologie requise pour rendre un tel projet réalisable : https://www.youtube.com/watch?v=2PEY0VV3ii0. La plupart des concepts que l'on trouve sur ces sites sont réalisables, au moins dans une certaine mesure, mais ce n'est pas encore pour aujourd'hui. Il s'agit d'un *vaporware* (ou fumiciel : un produit qu'on vous fait miroiter, qui n'existe pas encore mais qui est assez envisageable pour pouvoir attirer l'attention), mais l'idée est néanmoins intéressante.



Blue Origin, compagnie fondée par Jeff Bezos, dispose réellement d'une fusée fonctionnelle permettant de loger plusieurs personnes (https://www.csmonitor.com/Science/2017/0329/Blue-Origin-offers-window-into-what-a-space-vacation-might-look-like-literally). Cette fusée a déjà effectué à ce jour cinq voyages sans passagers. Il s'agirait d'emmener les gens non pas vraiment dans l'espace, mais plutôt à 100 km, en orbite terrestre basse. Des compagnies comme Blue Origin

(https://www.blueorigin.com/) et SpaceX (http://www.spacex.com/) ont actuellement les meilleures chances de faire des séjours dans l'espace une réalité. SpaceX est d'ailleurs en train d'étudier des projets de voyage sur Mars (http://www.spacex.com/mars).

Quel que soit ce que l'avenir nous réserve, des gens finiront par se rendre dans l'espace pour diverses raisons, entre autres pour leurs vacances. Certes, le coût du voyage sera tout aussi astronomique que la distance parcourue depuis la Terre. Les voyages dans l'espace ne risquent pas de devenir bon marché dans un avenir prévisible. Toujours est-il que des entreprises envisagent dès à présent des séjours dans le cosmos, mais qu'il faudra encore attendre avant d'avoir la possibilité d'embarquer.

Procéder à des investigations scientifiques

Beaucoup d'investigations scientifiques sont déjà menées dans l'espace, et l'IA y joue toujours un rôle. De la Station spatiale internationale au télescope Hubble, tout dépend fortement de l'IA (http://spacenews.com/beyond-hal-how-artificial-intelligence-is-changing-space-systems/). Concernant l'avenir, on peut prévoir des laboratoires tout équipés dans le cosmos, ou de courtes incursions dans l'espace pour y effectuer des expérimentations. La société Zero Gravity propose actuellement ce qu'elle appelle un vol parabolique à bord d'une « comète à vomi » pour vivre une expérience de quasi-apesanteur (https://www.gozerog.com/). En réalité, cette expérience a lieu dans un avion qui vole à haute altitude puis plonge en piqué. Cette tendance va sans doute se poursuivre, et à plus haute altitude.

Industrialiser l’espace

Il existe plusieurs manières de faire en sorte que les voyages dans l’espace deviennent rentables. Nous retirons déjà des avantages considérables des technologies développées pour les vols spatiaux et adaptées à une utilisation terrestre civile (un article parmi tant d’autres illustre l’importance de l’espace pour notre vie quotidienne sur terre : https://www.nasa.gov/press-release/spinoff-2016-highlights-space-technologies-used-in-daily-life-on-earth). Cependant, même avec les transferts de technologie, l’espace reste très cher, et l’on pourrait peut-être le rendre plus rentable en adaptant notre savoir-faire d’une autre manière, par exemple en créant des usines spatiales (https://www.popsci.com/factories-in-space).

Il se pourrait même que les usines spatiales soient le seul moyen d’obtenir certains matériaux et certains produits (à titre d’exemple, voir https://www.fastcodesign.com/3066988/mit-invented-the-material-well-need-to-build-in-space). L’absence de gravité change la manière dont les matériaux réagissent et s’associent, si bien que ce qui est impossible ici sur terre devient tout à fait possible dans l’espace. Par ailleurs, certains processus ne peuvent se réaliser facilement que dans l’espace, comme la fabrication de roulements à billes avec des billes parfaitement rondes (https://www.acorn-ind.co.uk/insight/The-Science-Experiment-Which-Took-Off-Like-A-Rocket---Creating-Space-Ball-Bearings/).

Utiliser l’espace pour le stockage

Nous finirons par stocker certains matériels dans l’espace, et ce sera judicieux. À mesure que les voyages dans l’espace deviendront de plus en plus fréquents et que commencera l’industrialisation de l’espace, le besoin d’y stocker des ressources telles que le carburant et les minerais extraits augmentera. Sachant que les débouchés des ressources minières ne seront pas toujours connus tout de suite (il faudra aussi des ressources pour les usines spatiales), il sera effectivement moins onéreux de les conserver dans l’espace plutôt que de les stocker sur Terre, jusqu’à ce que l’on en ait besoin. Les stations-service dans l’espace pourraient devenir une réalité plus rapidement qu’on le pense, sachant qu’elles se révéleront peut-être indispensables pour les voyages vers Mars.

Bien qu’il n’existe encore aucun projet de stockage de déchets dangereux dans l’espace, il se peut que cette option se matérialise un jour, comme moyen d’éviter que ces déchets polluent notre planète. Une question qui vient à l’esprit, naturellement, est de savoir pourquoi il conviendrait de stocker des déchets dangereux quelque part plutôt que les envoyer brûler dans le soleil, par exemple. Sur cette lancée, on peut aussi se demander s’il est indispensable de continuer à produire des déchets dangereux. Or, tant qu’il y aura des humains, il y aura production de déchets dangereux. Stocker ces déchets dans l’espace nous laisserait une chance de trouver un moyen de les recycler et d’en faire quelque chose d’utile, tout en les maintenant à l’écart.

Chapitre 17
Inaugurer de nouvelles activités humaines

DANS CE CHAPITRE

» **Gagner de l'argent dans l'espace**

» **Construire des villes dans des lieux inédits**

» **Renforcer les capacités humaines**

» **Réparer notre planète**

Quand on entend parler de nouvelles avancées dans le domaine des robots, de l'automatisation et de l'IA, on a tendance à en voir les aspects négatifs plus facilement que les aspects positifs. Ainsi, par exemple, l'automatisation devrait faire disparaître entre 400 et 800 millions d'emplois d'ici 2030 si l'on en croit cet article : https://www.theverge.com/2017/11/30/16719092/automation-robots-jobs-global-800-millionforecast. L'article explique aussi comment ces emplois vont disparaître. Même si l'auteur admet que certaines avancées technologiques créent des emplois (on estime, par exemple, que l'ordinateur individuel en a créé 18,5 millions), l'accent est mis sur les emplois perdus et sur le risque que cette perte soit permanente (comme on est censé l'avoir observé dans le secteur de l'industrie). Le problème est que les auteurs de ce genre d'article ont

tendance à être catégoriques à propos des pertes d'emplois, mais à rester dans le vague, dans le meilleur des cas, lorsqu'il s'agit de création d'emplois. L'objectif global de ce chapitre est d'en finir avec tout ce battage, cette désinformation et cet alarmisme, et de mettre l'accent sur de meilleures nouvelles.

Ce chapitre est consacré à de nouvelles activités humaines intéressantes. Cependant, pour commencer, ne pensez pas que votre emploi est menacé (pour quelques exemples d'activités que l'IA ne pourra pas menacer, voir Chapitre 18). À moins que votre activité soit abrutissante de simplicité et extrêmement répétitive, vous ne risquez pas d'être remplacé par une IA. Tout au contraire, vous vous apercevrez qu'une IA vous apporte un « plus » et vous permet de tirer davantage de satisfaction de votre métier. Cependant, quand vous aurez lu ce chapitre, peut-être aurez-vous envie de bénéficier d'une formation pour vous lancer dans une activité vraiment nouvelle et formidable.



Certaines activités mentionnées dans ce chapitre sont aussi quelque peu dangereuses. L'IA ajoutera aussi des applications triviales à la liste des tâches que vous devrez effectuer au bureau, ou peut-être même chez vous, mais ce seront les éléments les plus intéressants de cette liste, et cette nouvelle activité reste attrayante même si une IA vous remplace dans ce que vous faisiez jusqu'ici. Surtout, cette situation s'est reproduite à maintes reprises dans l'histoire, de la façon la plus perturbatrice au moment de la Révolution industrielle, mais nous avons toujours réussi à trouver de nouvelles occupations. Si vous ne deviez retenir qu'une seule chose de ce chapitre, c'est que tout cet alarmisme n'a pas de fondement réel : simplement, quelqu'un essaie de vous faire peur et de vous faire croire une chose qui n'est pas vraie.

Habiter et travailler dans l’espace

Les médias ont mis dans la tête des gens l’idée que nous devrons un jour parcourir l’univers ou mener de grandes batailles dans l’espace contre des extraterrestres venus conquérir la Terre. Le problème est que la plupart des gens ne sauraient faire ni l’un ni l’autre. Or, il est tout à fait possible aujourd’hui d’exercer un métier dans lequel certaines tâches sont à finalité spatiale, par exemple dans une société comme SpaceX (voir http://www.spacex.com/careers). Les possibilités d’emplois sont même considérables (http://www.spacex.com/careers/list), et un certain nombre de ces emplois sont des stages, si bien que vous pouvez acquérir de l’expérience avant de faire le grand saut. Bien sûr, il s’agit d’activités à caractère technique, mais vous pouvez voir qu’il y a dans cette liste un peu de tout. Au moment où ces lignes sont écrites, cette liste comprend même un emploi de garçon de café. En fait, les carrières spatiales comprendront tout ce que comprennent les autres carrières : simplement, vous aurez finalement la possibilité d’évoluer vers quelque chose de plus intéressant.



Des entreprises comme SpaceX prévoient aussi leurs propres possibilités de formation et d’échanges avec les universités (http://www.spacex.com/university). L’espace est une aventure relativement nouvelle, si bien que tout le monde commence à peu près au même niveau, dans la mesure où tout le monde se forme à quelque chose de nouveau. Ce qui est passionnant quand on aborde un nouveau domaine de l’aventure humaine, c’est de commencer à faire ce qui n’a jamais été fait auparavant. Par conséquent, il y a une courbe d’apprentissage. Vous pourriez trouver là une possibilité d’apporter une contribution vraiment significative à l’humanité, mais seulement si vous acceptez de relever le

défi qui consiste à découvrir quelque chose de différent et à assumer les risques que cela suppose.

Pour l'instant, les possibilités de vivre et de travailler dans l'espace sont limitées, mais elles se développeront avec le temps. Le Chapitre 16 mentionne toutes sortes de choses que les humains finiront par faire dans l'espace, comme l'exploitation minière ou la recherche. Oui, quand nous aurons visité d'autres planètes, nous finirons pas fonder des cités dans l'espace. Mars pourrait devenir la prochaine Terre. La planète Mars a souvent été considérée comme potentiellement habitable (voir par exemple http://www.planetary.org/blogs/guest-blogs/2017/20170921-mars-isru-tech.html et https://www.nasa.gov/feature/goddard/2017/mars-mission-sheds-light-on-habitability-of-distant-planets), à condition toutefois d'y recréer la magnétosphère (https://phys.org/news/2017-03-nasa-magnetic-shield-mars-atmosphere.html).

Concernant la vie dans l'espace, certaines idées dont on entend parler actuellement ne semblent pas réalisables, mais on y pense très sérieusement et ce sont des choses théoriquement possibles. Ainsi, une fois que l'on aurait restauré la magnétosphère de la planète Mars, il serait possible d'y reconstituer une biosphère de manière à la rendre tout à fait habitable (il existe de nombreux articles sur ce sujet, et celui-ci explique comment il serait possible de rendre l'environnement de Mars suffisamment oxygéné. Certains de ces changements se réaliseraient automatiquement, tandis que d'autres nécessiteraient une intervention humaine. Imaginez ce que ce serait, de faire partie d'une équipe chargée de créer une biosphère. Cela dit, la viabilité de ce genre de projet dépendra beaucoup des IA, qui sont capables de voir ce que nous ne voyons pas et de produire des réponses que nous ne pouvons même pas imaginer aujourd'hui. Les humains et les IA

travailleront ensemble à remodeler des endroits comme Mars pour répondre aux besoins de l'humanité. Plus important, ces efforts feront appel à un nombre considérable d'individus ici, sur Terre, ainsi que sur la Lune, dans l'espace et sur Mars. Une bonne coordination sera essentielle.

Créer des villes dans des environnements hostiles

Au moment où ces lignes sont écrites, la Terre compte 7,6 milliards d'habitants (http://www.worldometers.info/world-population/), et ce nombre ne fera qu'augmenter. Au cours d'une seule journée, la population aura augmenté de 153030 individus. En 2030, date à laquelle la NASA prévoit de tenter la première expédition sur Mars, la Terre comptera 8,5 milliards d'habitants. La population terrestre est donc très nombreuse, et elle le sera encore davantage demain. À un moment donné, il sera nécessaire de trouver d'autres endroits pour vivre. Si une chose est sûre, c'est qu'il faudra trouver de nouvelles surfaces pour l'agriculture. En même temps, nous souhaitons conserver des espaces naturels dans le monde, et consacrer aussi des surfaces à d'autres finalités. Heureusement, l'IA peut nous permettre d'identifier des endroits appropriés pour construire, de découvrir comment rendre viable le processus de construction et de conserver un environnement adapté une fois qu'un nouveau lieu sera disponible.

À mesure que l'IA et les humains acquièrent de nouvelles aptitudes, des milieux hostiles deviennent plus accessibles. En théorie, nous devrions pouvoir un jour construire des habitats dans un volcan, mais en attendant, il existe évidemment des endroits plus propices. Les

sections qui suivent mentionnent quelques endroits qui pourraient un jour convenir pour y créer des villes. Ils offrent tous des avantages que nous n'avions jamais eus jusqu'ici : des possibilités pour nous de faire progresser nos connaissances et notre aptitude à vivre dans des milieux plus hostiles encore dans l'avenir.

Construire des villes dans l'océan

Il existe plusieurs manières de construire des villes dans l'océan, mais les deux plus répandues sont les villes flottantes et les villes bâties sur le plancher océanique. Une ville flottante est actuellement au stade de projet au large des côtes de Tahiti (http://www.dailymail.co.uk/sciencetech/article-4127954/Plans-worlds-floating-city-unveiled.html). Les villes flottantes répondent à divers objectifs, et plus particulièrement :

» la protection contre la hausse du niveau des mers ;

» des possibilités d'expérimenter de nouvelles méthodes agricoles ;

» le développement de nouvelles techniques de pisciculture ;

» la création d'une nouvelle forme de gouvernement.

La colonisation humaine sur la mer est ce que l'on appelle dans le milieu anglophone le *seasteading*. Les premières villes existeront dans des zones relativement protégées. Il est tout à fait possible de construire sur l'océan même, toutefois cela revient cher (les plateformes d'exploitation pétrolière utilisent déjà différentes sortes d'IA pour rester stables et exécuter des tâches différentes : voir

https://www.techemergence.com/artificial-intelligence-in-oil-and-gas/).

Les villes sous l'eau sont aussi parfaitement envisageables, et il existe déjà des laboratoires de recherche sous-marins. Aucun de ces laboratoires ne se trouve dans les grandes profondeurs, mais même une profondeur d'une vingtaine de mètres sous l'eau est déjà assez appréciable. D'après un certain nombre de sources, la technologie permettant de construire des villes de plus grande dimension et plus en profondeur existe, mais un meilleur suivi serait nécessaire. C'est là que l'IA devrait entrer en jeu. L'IA pourrait assurer la surveillance de la ville depuis la surface et contrôler les dispositifs de sécurité qu'une telle ville exigerait.



Il est important de considérer que ces villes dans l'océan ne ressembleront sans doute pas aux villes terrestres. Ainsi, par exemple, des architectes envisagent de construire près de Tokyo une ville sous-marine qui ressemblerait à une spirale géante (http://www.businessinsider.com/underwater-city-tokyo-japan-2017-1). Cette spirale pourrait héberger jusqu'à 5000 personnes. Cette ville très particulière serait située à 5400 mètres sous la surface et sa viabilité reposerait sur des technologies avancées, notamment en matière d'énergie. Ce serait une ville à part entière, avec par exemple des laboratoires, des restaurants et des écoles.

Quelle que soit la manière dont les gens iraient s'y installer, l'IA serait mise à contribution de façon généralisée. Cette IA est déjà partiellement en cours d'élaboration (http://news.mit.edu/2017/unlocking-marine-mysteries-artificial-intelligence-1215), et des étudiants travaillent avec des robots subaquatiques. Comme on peut l'imaginer, les robots contribueront au

développement urbain sous-marin. Ils effectueront diverses tâches de maintenance qui seraient impossibles pour des êtres humains.

Créer des habitats dans l’espace

Un *habitat spatial*, contrairement aux autres formes de station dans l’espace, est une colonie permanente. L’intérêt d’un habitat spatial est de fournir à des personnes un hébergement de longue durée. Il s’agira sans doute d’un milieu en vase clos, dans lequel les gens pourront vivre sans devoir s’approvisionner indéfiniment à l’extérieur (ou peu s’en faudra). Par conséquent, il faudra que l’air et l’eau y soient recyclés, que l’on puisse y faire des cultures aux fins alimentaires, et que les habitants disposent de moyens d’effectuer d’autres tâches que n’offrent pas les stations spatiales dans lesquelles le séjour ne peut être que de courte durée. Si toutes les stations spatiales nécessitent une IA pour surveiller et optimiser les conditions de vie, l’IA d’un habitat spatial sera considérablement plus complexe.

Le Chapitre 16 aborde le sujet des habitats dans l’espace dans la section « Passer ses premières vacances dans l’espace ». Naturellement, les premiers contacts du public avec l’espace prendront la forme de visites brèves. Des vacances dans l’espace, ce sera sûrement très intéressant ! Néanmoins, un séjour en orbite terrestre et un habitat à long terme dans l’espace lointain sont deux choses bien différentes, et la seconde s’imposera pour la NASA si elle réussit effectivement à faire du voyage jusqu’à la planète Mars une réalité. La NASA a déjà chargé six compagnies de commencer à étudier les conditions requises pour créer des habitats dans l’espace profond. Vous pouvez voir des prototypes fabriqués par

ces compagnies sur cette page : https://www.nasa.gov/feature/nextstep-partnerships-develop-ground-prototypes.

Pour certaines organisations, les habitats dans l'espace ne sont pas tant un moyen de développer l'exploration que de protéger la civilisation. Si un astéroïde géant devait un jour percuter la Terre, la plus grande partie de l'humanité périrait, mais les habitants de la Station spatiale internationale (SSI) pourraient survivre, du moins, si l'astéroïde ne l'a pas frappée également. Cependant, la SSI n'est pas viable comme stratégie de survie à long terme, et le nombre de personnes qu'elle peut accueillir à tout moment est limité. C'est pourquoi la Lifeboat Foundation, par exemple (https://lifeboat.com/ex/spacehabitats), étudie la possibilité d'aménager des habitats dans l'espace afin d'assurer la survie de l'humanité. Sa première tentative se nomme Ark I (https://lifeboat.com/ex/arki), un projet conçu pour héberger 1000 résidents permanents et jusqu'à 500 invités. Théoriquement, la technologie est viable, mais une importante planification est nécessaire.

Un autre emploi des habitats spatiaux est celui de *vaisseau générationnel*, un type de vaisseau servant à explorer l'espace interstellaire à l'aide de technologies dont nous disposons aujourd'hui. Les gens vivraient dans ce vaisseau durant son voyage vers les étoiles. Ils auraient des enfants dans l'espace, et c'est ce qui permettrait d'envisager des voyages aussi longs. L'idée du vaisseau générationnel n'est pas nouvelle. On la retrouve depuis longtemps dans la littérature de science-fiction et dans le cinéma fantastique. Il existe cependant des tentatives de créer réellement un vaisseau générationnel : *http://www.icarusinterstellar.org/building-blocks-for-a-generation-ship* . Le problème est qu'il faudrait disposer d'un nombre constant de personnes prêtes à travailler

dans chacune des spécialités indispensables au fonctionnement du vaisseau. Même dans ce cas, grandir en sachant qu'un travail essentiel vous attend représenterait un changement intéressant par rapport à notre condition présente.



Plutôt que de construire sur terre les éléments d'un habitat spatial pour les envoyer ensuite dans l'espace, la stratégie actuelle consiste à extraire des astéroïdes les matériaux nécessaires et à utiliser des usines spatiales pour produire les habitats. On estime que la principale ceinture d'astéroïdes du système solaire contient des matériaux en quantités suffisantes pour construire des habitats représentant autant de surface que 3000 fois la Terre : de quoi héberger une sacrée quantité d'êtres humains dans l'espace !

Créer des ressources sur la Lune

La question n'est pas de savoir si nous allons retourner sur la Lune pour y installer des bases, mais quand nous allons le faire. Les stratégies actuelles en matière de colonisation de l'espace dépendent souvent de ressources lunaires de différentes sortes, et c'est notamment le cas du projet de la NASA d'envoyer des êtres humains sur Mars. Les projets de bases lunaires ne manquent pas non plus. Vous pouvez en découvrir quelques-uns ici : *https://interestingengineering.com/8-interesting-moon-base-proposals-every-space-enthusiast-should-see*



À un moment donné, il a été question de bases militaires sur la Lune (http://www.todayifoundout.com/index.php/2017/01/project-horizon/), mais le Traité de l'espace, signé par 60 pays dans le but d'éviter la politisation de

l’espace (http://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/introouterspacetreaty.html), a largement mis fin à cette idée. Les structures lunaires et les services qu’elles assureront répondront plus vraisemblablement, dans un premier temps, à des besoins en matière d’exploration, d’extraction minière et de fabrication, puis viendra la construction de villes entières. Ces projets feront sans doute intervenir des robots, mais les humains resteront indispensables pour tout un ensemble de tâches, notamment la réparation et la gestion des robots. La construction de bases sur la Lune requerra aussi tout un ensemble de nouvelles activités qui ne seront probablement pas requises dans les projets d’habitats ni dans les situations qui concerneront exclusivement les travaux dans l’espace. Il faudra, par exemple, que quelqu’un s’occupe des conséquences des séismes lunaires (voir https://science.nasa.gov/science-news/science-at-nasa/2006/15mar_moonquakes).

Exploiter les caractéristiques du sol lunaire pour construire des logements est aussi une possibilité. La découverte récente de structures répondant aux besoins de la colonisation peut faciliter la construction de bases lunaires. Il est question, par exemple, d’une gigantesque grotte qui conviendrait pour la colonisation : http://time.com/4990676/moon-cave-base-lunarcolony-exploration/. En l’occurrence, le Japon a découvert ce qui semble être un tube de lave qui protégerait les colons de diverses menaces environnementales.



Naturellement, les exagérations dans les médias à propos de ces structures (qui sont très vraisemblablement d’origine naturelle) sont stupéfiantes. À en croire certaines sources, des structures sur la face cachée de la Lune auraient été fabriquées par des extraterrestres

(http://www.dailymail.co.uk/sciencetech/article-4308270/UFO-hunters-claim-footage-aliens-moon.html). Les structures existent : nous pouvons les utiliser pour faciliter la construction de bases. À vous de faire le tri entre les sources d'information.

Rendre les humains plus efficients

Une IA peut rendre l'humain plus efficient de différentes façons. La plupart des chapitres de ce livre présentent un exemple dans lequel un être humain s'aide d'une IA pour faire les choses de manière plus efficiente. Cependant, un des exemples les plus intéressants est celui du Chapitre 7, qui indique de quelle manière une IA permettra de répondre aux besoins des gens dans le domaine médical. Dans toutes ces utilisations d'une IA, on suppose qu'un être humain reste en commande, mais qu'il utilise l'IA pour devenir meilleur dans son activité. Le système chirurgical da Vinci, par exemple, ne remplace pas le chirurgien : il le rend simplement capable d'accomplir sa tâche avec plus de facilité et moins de risque d'erreur. Une nouvelle activité qui accompagne ce progrès consiste à apprendre aux professionnels à utiliser de nouveaux outils qui intègrent une IA.

LES HABITATS OU LA « BIOSPHÉRISATION »

L'IA est appelée à être beaucoup utilisée, quelles que soient les décisions qui seront prises en ce qui concerne l'habitat et les activités dans l'espace. L'orientation donnée au développement de l'IA sera différente selon les destinations et le moment choisi. Actuellement, l'idée fait son chemin que des hommes pourraient vivre sur Mars dans un délai relativement court. Or, quand on se rend sur des sites Internet comme *https://phys.org/news/2017-03-future-space-colonization-terraforming-habitats.html*, il devient évident que la « biosphérisation » de Mars demandera beaucoup de temps. Rien que le réchauffement de la planète (après avoir installé la technologie nécessaire pour y recréer la magnétosphère) demandera environ un siècle. Par conséquent, entre les habitats et la « biosphérisation », nous n'avons pas le choix : ce seront d'abord les habitats, et ils seront sans doute largement utilisés pour les projets éventuels d'activités sur Mars. L'IA n'en sera pas moins différente pour les deux projets, et il sera intéressant de voir quels seront les problèmes que l'IA permettra de résoudre.



Dans l'avenir, il existera des consultants dont la seule tâche consistera à trouver de nouveaux moyens d'intégrer des IA dans les processus opérationnels pour rendre les gens plus efficients. Dans une certaine mesure, cette profession existe déjà, mais il s'agit d'un besoin qui augmentera à un moment donné, quand les IA génériques et configurables deviendront monnaie courante. Pour beaucoup d'entreprises, la rentabilité sera assurée en trouvant l'IA qui augmentera les travailleurs humains en leur permettant d'accomplir leurs tâches sans erreur et aussi rapidement que possible. Leur rôle sera en partie celui d'analystes-programmeurs, en partie celui de commerciaux et en partie celui de formateurs, tout cela

rassemblé dans une seule fonction. Un exemple de cette conception des profils futurs est présenté dans cet article : http://www.information-age.com/harness-ai-improve-workplace-efficiency-123469118/.

En termes d'efficience humaine, songeons aux domaines dans lesquels une IA peut exceller. Par exemple, une IA ne fera pas bien l'affaire dans une tâche créative : la créativité doit donc rester le domaine de l'humain. En revanche, l'IA est très performante lorsqu'il s'agit de chercher, par conséquent ce genre de tâche pourra être confiée à une IA tandis que les humains se consacreront à des tâches plus créatives. Voici quelques possibilités futures d'utilisation d'une IA pour devenir plus efficient :

» **L'embauche :** actuellement, les recruteurs ne connaissent pas toujours très bien le cursus et les aptitudes réelles des candidats qu'ils reçoivent. Une IA pourrait effectuer des recherches sur le candidat avant l'entretien d'embauche, afin que le recruteur dispose de plus d'informations. En outre, sachant que l'IA utiliserait la même méthodologie de recherche pour tous les candidats, l'entreprise aurait la garantie qu'ils seraient tous traités de façon juste et équitable.

» Le géant du secteur des biens de consommation Unilever utilise aussi ce type de technologie.

» **L'ordonnancement :** aujourd'hui, une entreprise est sans cesse confrontée à un risque lorsque quelqu'un a négligé de programmer une tâche donnée. Cette personne n'a peut-être pas trouvé le temps de le faire, ou bien elle n'a même pas pensé que cette tâche allait être nécessaire. Les secrétaires et les assistants avaient l'habitude de gérer les horaires, mais dans les nouvelles hiérarchies réduites ou horizontales, ces assistants ont disparu et chacun s'occupe de planifier ses

propres tâches. Ainsi, les salariés manquent souvent des occasions de permettre à leur entreprise d’exceller parce qu’ils sont trop occupés à gérer leur emploi du temps. Quand une IA est couplée avec une personne, celle-ci n’a plus besoin de se charger de la programmation des tâches. Elle peut alors voir plus loin et prévoir ce qui devra être programmé. Le salarié apporte et rapporte davantage à l’entreprise quand il peut se consacrer aux activités dans lesquelles il excelle parce qu’il n’a plus à s’occuper du reste. L’IA favorise donc grandement l’excellence humaine.

» **Trouver les informations cachées :** aujourd’hui, plus que jamais, les entreprises sont prises en défaut par la concurrence en raison d’informations cachées. La surcharge d’informations et la complexité croissante de la science, des technologies, des entreprises et de la société sont à la racine du problème. Peut-être existe-t-il une nouvelle façon de conditionner les produits, qui réduirait significativement les coûts, ou peut-être la structure de l’entreprise change-t-elle par suite de la politique interne. Ce n’est qu’en sachant à tout moment ce qui est disponible et ce qui est en train de se passer qu’une entreprise peut véritablement prospérer, mais une telle situation n’est pas envisageable. Si l’on s’accordait le temps nécessaire pour savoir tout ce qu’il y a à savoir dans une activité particulière, il ne resterait plus de temps pour exercer cette activité.

» Or, l’IA est extraordinairement performante lorsqu’il s’agit de trouver les choses. Grâce au recours à l’apprentissage machine, on pourrait apprendre à une IA à identifier avec précision les bonnes questions et les exigences pour maintenir l’entreprise à flot sans devoir gaspiller tant de temps à faire des recherches manuellement.

» **L'aide adaptive** : actuellement, quiconque utilise des produits logiciels doit bien admettre qu'il est parfois très frustrant de devoir se rappeler comment une certaine tâche doit être réalisée, surtout quand il faut pour cela recourir à l'aide de l'application. On peut déjà voir comment une IA devient une aide adaptive lorsqu'il s'agit de saisir certaines formes d'informations. Cependant, une IA peut faire bien davantage. En utilisant les techniques de l'apprentissage machine pour identifier des habitudes d'utilisation, une IA pourrait fournir aux utilisateurs une aide adaptive qui leur permettrait de ne plus buter sur les parties de l'application dont il est difficile de se rappeler comment elles sont conçues. Parce que chaque utilisateur est différent, une application programmée pour fournir une aide adaptive ne fonctionnera jamais. Le recours à l'apprentissage machine permet de personnaliser le système d'aide pour l'adapter à chaque utilisateur.

» **L'apprentissage adaptatif** : il vous est aujourd'hui possible de passer un test adaptatif qui s'adapte lui-même à vous et vous pose des questions sur les points faibles détectés dans vos connaissances. Le test adaptatif constate que vous en savez suffisamment, ou bien, il vous pose suffisamment de questions pour vérifier que vous avez besoin d'une formation. Un jour, les applications seront capables de détecter la manière dont vous les utilisez et de vous dispenser automatiquement une formation qui vous rendra plus performant. Une application pourra déterminer, par exemple, que vous pourriez exécuter une tâche avec moins de clics, et vous montrer comment. En apprenant constamment aux gens à utiliser la méthode la plus efficiente, notamment quand ils travaillent à l'ordinateur, on leur permet de

> progresser, mais le rôle de l’être humain reste indispensable.

Résoudre les problèmes à l’échelle planétaire

Que l’on croie ou non au réchauffement planétaire, que l’on pense ou non que la pollution est un problème et que l’on se soucie ou non de la surpopulation, le fait est que nous n’avons qu’une planète et que la Terre est confrontée à des problèmes. À coup sûr, le climat se dérègle, de vastes régions ne sont plus exploitables à cause de la pollution, et certaines régions du monde sont indubitablement surpeuplées. Les tempêtes et les feux de forêt incontrôlables se moquent de ce que vous pensez, et le résultat est toujours le même : la destruction de zones habitées. Quand trop de gens sont entassés dans un espace trop limité, il en résulte généralement des épidémies, de la délinquance et de la criminalité, et d’autres problèmes encore. Ce n’est pas une question d’idées politiques ni de croyance personnelle. Il s’agit de problèmes réels, et l’IA peut les résoudre en indiquant aux personnes compétentes quelles sont les lignes à suivre. Les sections qui suivent abordent les problèmes planétaires du point de vue de l’utilisation de l’IA pour les identifier, les comprendre, et éventuellement les résoudre. Nous ne cherchons pas à faire passer ici un quelconque message de nature politique ou autre.

Réfléchir à la façon dont le monde fonctionne

Aujourd'hui, les capteurs permettent de surveiller tous les aspects de la planète. Les informations sont même si abondantes qu'on peut admirer cette possibilité que nous avons de les rassembler toutes, tandis qu'il est bien plus difficile de les exploiter. Par ailleurs, en raison des interactions entre différents milieux terrestres, on ne peut pas réellement savoir quels faits ont des répercussions sur d'autres parties de l'environnement. Ainsi, il est difficile de savoir précisément dans quelle mesure les vents ont un effet sur le réchauffement des mers, lequel influence les courants qui sont susceptibles d'engendrer des tempêtes. Si nous comprenions vraiment toutes ces interactions, le bulletin météorologique pourrait être plus exact. Malheureusement, il est assez exact en général, du moins, si l'on n'est pas trop regardant. Si nous acceptons ce niveau de performance chez les météorologues, cela montre que nous sommes conscients de la difficulté de leur tâche.

Avec le temps, les prévisions météorologiques sont devenues bien plus fiables, et c'est en partie grâce à tous ces capteurs installés un peu partout. On a aussi mis au point de meilleurs modèles météorologiques et constitué de bien plus vastes magasins de données aux fins de prévisions climatiques. Néanmoins, la principale raison de ce progrès dans la précision est le recours à l'IA pour éplucher les chiffres et repérer des tendances dans les données résultantes (voir https://www.techemergence.com/ai-for-weather-forecasting/).

La météo est en réalité un des phénomènes terrestres les mieux compris. Songeons à la difficulté de prévoir les séismes. Grâce à l'apprentissage machine, les scientifiques devraient plus facilement savoir à quel moment un tremblement de terre va se produire (https://www.express.co.uk/news/science/871022/earthquake-artificialintelligence-AI-

cambridge-university), mais seul le temps dira si ces nouvelles informations sont réellement utiles. Il fut un temps où l'on pensait que les conditions météorologiques avaient une incidence sur les séismes, mais ce n'est pas le cas. À l'inverse, les séismes peuvent avoir une incidence sur la météo, en changeant les conditions environnementales. Les séismes et la météo peuvent aussi, ensemble, aggraver encore la situation (https://www.usatoday.com/story/news/nation/2015/05/02/kostigen-earthquakeweather/26649071/).

Ce qui est encore plus difficile à prédire, ce sont les éruptions volcaniques. Au moins, la NASA est maintenant capable de détecter et d'obtenir des images des éruptions avec une remarquable précision.

Les éruptions volcaniques provoquent souvent des séismes, si bien qu'anticiper les premières permet de prédire ces derniers (http://volcano.oregonstate.edu/how-arevolcanoes-and-earthquakes-related). Naturellement, les volcans ont aussi une incidence sur la météo.

Les phénomènes naturels abordés jusqu'ici dans cette section ne sont que la partie émergée de l'iceberg. Si vous pensez que la complexité de la Terre est telle que personne ne pourra jamais tout comprendre, vous voyez juste. C'est pourquoi nous avons besoin de créer des IA et de leur apprendre à aider les humains à mieux comprendre comment le monde est fait. Le développement de ce type de connaissance pourrait permettre d'éviter des catastrophes dans l'avenir, et de limiter les effets de certains maux d'origine humaine.



Quoi que vous ayez pu lire ici ou là, il n'existe à l'heure actuelle aucun moyen d'éviter le mauvais temps, les tremblements de terre ou les éruptions volcaniques. Le

mieux que l'on puisse espérer réussir aujourd'hui, c'est prédire ces événements puis agir pour en limiter l'impact. Néanmoins, cette capacité de réduire l'impact des phénomènes naturels représente à elle seule un grand pas en avant. Avant l'IA, nous étions à la merci des événements car il était impossible de prédire une catastrophe naturelle avant qu'il soit trop tard pour vraiment agir de façon proactive afin d'en limiter les conséquences.

De même, même s'il peut sembler possible de prévenir toutes les catastrophes d'origine humaine, ce n'est souvent pas le cas. Même avec la meilleure planification possible, on ne pourra jamais empêcher que des accidents surviennent. Cela dit, la majorité des effets des activités humaines peuvent être contrôlés et éventuellement évités si l'on dispose des connaissances adéquates, que l'on peut obtenir grâce au filtrage par motifs effectué par une IA.

Identifier les sources potentielles de problèmes

Avec tous ces yeux que nous avons maintenant dans le ciel, on peut penser que les satellites sont une source ultime de données permettant de prédire les phénomènes terrestres. Il convient cependant de ne pas négliger un certain nombre de problèmes :

» La Terre est immense, si bien que pour détecter un phénomène particulier, il est nécessaire de passer au crible plusieurs millions d'images à chaque seconde de la journée.

» Les images doivent apparaître avec une résolution suffisamment élevée pour qu'un phénomène puisse vraiment être identifié.

» Il est essentiel d’utiliser le filtre de lumière adéquat, car certains phénomènes ne deviennent visibles qu’à travers un certain filtre.

» Les conditions météorologiques peuvent empêcher l’acquisition de certains types d’images.

Malgré tous ces problèmes, les chercheurs et autres utilisateurs recourent à l’IA pour parcourir les photographies prises chaque jour et chercher d’éventuelles anomalies (https://www.cnet.com/news/descartes-labs-satellite-imagery-artificial-intelligence-geovisual-search/). Cependant, l’IA ne peut indiquer d’éventuels problèmes et effectuer une analyse que si les images apparaissent sous un format adéquat. Il faut encore que quelqu’un détermine si le problème est réel et s’il y a lieu de le résoudre. Ainsi, par exemple, une forte tempête au milieu de l’océan Pacifique, loin des terres et des voies maritimes, ne sera sans doute pas considérée comme un problème prioritaire. En revanche, une tempête similaire au-dessus des terres sera une source de préoccupations. Naturellement, il est toujours préférable de détecter une tempête avant qu’elle devienne un problème, plutôt que d’essayer d’agir une fois qu’elle a sévi.



Outre le balayage des images pour identifier d’éventuels problèmes, l’IA peut aussi améliorer les images. L’article de la page https://www.wired.com/story/how-ai-could-really-enhance-images-from-space/ explique comment une IA peut améliorer la résolution des photos prises depuis l’espace et les rendre plus exploitables. En améliorant ces images, l’IA peut mieux déterminer certains types particuliers de phénomènes d’après leurs propriétés typiques. Bien sûr, pour pouvoir formuler une prédiction, il faut que l’IA ait déjà eu connaissance de schémas similaires à celui qu’elle va

détecter. Les humains auront toujours besoin de vérifier ce que fait l'IA et de vérifier que le phénomène identifié correspond effectivement à ce que l'IA caractérise.

Définir des solutions potentielles

La solution à un problème planétaire dépend de ce problème. S'il s'agit d'une tempête, par exemple, ou d'un séisme, ou d'une éruption volcanique, il n'est même pas question d'essayer de prévenir le phénomène. Le mieux que l'on puisse espérer réussir, à l'heure actuelle, c'est d'évacuer la zone concernée et de reloger les habitants quelque part. Néanmoins, en en sachant le plus possible et le plus tôt possible, on peut répondre de façon proactive plutôt qu'intervenir seulement une fois que le chaos est là.

D'autres phénomènes ne nécessitent pas toujours une évacuation. Avec les techniques actuelles et un peu de chance, on peut limiter l'incidence d'un feu de forêt, par exemple. Déjà, il arrive que les professionnels des services d'incendie utilisent l'IA pour prédire les feux de forêt avant qu'ils ne se déclarent (https://www.ctvnews.ca/sci-tech/artificialintelligence-can-better-predict-forest-fires-says-albertaresearcher-1.3542249). Si l'utilisation de l'IA pour pouvoir identifier un problème puis élaborer une solution sur la base des données historiques est envisageable, c'est parce que dans le passé, les humains ont enregistré toutes ces informations sur ces phénomènes.

Quand on étudie les problèmes planétaires, l'exploitation des données historiques est essentielle. Pour résoudre un problème, se contenter d'une seule solution potentielle

n’est généralement pas une bonne idée. Mieux vaut en envisager plusieurs, et une IA peut permettre de les ordonner en fonction de l’historique des résultats. Bien sûr, là encore, un être humain distinguera dans les solutions quelque chose qui fera qu’une solution sera préférable à une autre. Ainsi, par exemple, il arrive qu’une solution particulière ne soit pas viable parce que les ressources qu’elle implique ne sont pas disponibles, ou parce que les intervenants n’ont pas eu la formation adéquate.

Voir les effets des solutions

Pour assurer le suivi des résultats d’une solution particulière, il faut enregistrer les données en temps réel, les analyser aussi rapidement que possible, puis en montrer les effets d’une manière qui soit compréhensible pour les humains. Une IA peut rassembler les données, les analyser et produire plusieurs présentations de ces données bien plus rapidement qu’un humain ne pourra jamais le faire. Nous n’avons pas fini d’établir les critères pour mener à bien toutes ces tâches et prendre les décisions finales : l’IA intervient simplement comme un outil permettant à l’utilisateur humain d’agir dans un délai raisonnable.



Dans l’avenir, certains pourraient se spécialiser dans les interactions avec l’IA en vue d’améliorer le traitement des données. Souvent, pour obtenir les bons résultats, il faut savoir quelle question poser et comment la poser. Aujourd’hui, si l’IA donne souvent de mauvais résultats, c’est parce que les utilisateurs ne sont pas assez familiarisés avec son fonctionnement pour pouvoir lui poser des questions pertinentes.

Ceux qui supposent qu’une IA pense de la même manière qu’un être humain ne pourront pas en tirer de bons résultats. Bien sûr, c’est ce qui est actuellement mis en

avant dans notre société. Dans la publicité pour Siri et pour Alexa, l'IA apparaît comme quasi humaine, ce qu'elle n'est évidemment pas. Or, même si c'est à une IA qu'ils s'adressent, des utilisateurs confrontés à une situation d'urgence doivent savoir lui poser des questions appropriées et les poser d'une manière adaptée s'ils veulent en tirer les résultats nécessaires. Vous ne pouvez pas voir les effets d'une solution si vous ne savez pas ce que vous pouvez attendre de l'IA.

Essayer encore

La Terre est un endroit compliqué. Divers facteurs interagissent les uns avec les autres, d'une manière que personne ne peut prévoir. Par conséquent, la solution que vous avez élaborée ne vous permettra peut-être pas de résoudre le problème. Si vous suivez l'actualité, vous aurez sans doute remarqué que bien des solutions ne résolvent rien du tout. C'est en procédant par tâtonnements que l'on peut comprendre ce qui est efficace et ce qui ne l'est pas. Néanmoins, en faisant appel à une IA pour identifier les modèles d'échec – les solutions qui n'ont pas fonctionné, et la raison à cela – vous pouvez limiter le nombre de solutions que vous aurez besoin d'essayer pour trouver laquelle est efficace. Par ailleurs, une IA peut étudier des scénarios similaires pour des solutions qui ont été probantes dans le passé, ce qui permet parfois d'économiser du temps et des efforts quand on cherche de nouvelles solutions à tester. L'IA n'est pas une baguette magique qu'il suffirait d'agiter pour créer une solution qui fonctionnerait tout de suite. La raison pour laquelle les humains auront toujours un rôle à jouer est que seuls les humains peuvent voir les résultats pour ce qu'ils sont.

Actuellement, une IA est toujours programmée pour gagner. L'encadré « Comprendre l'orientation de



l'apprentissage » du Chapitre 13 examine la possibilité de créer une IA qui comprenne la futilité, c'est-à-dire la situation dans laquelle il n'y a pas de gagnant possible. Or, une telle IA n'existe pas encore et n'existera peut-être jamais. Les humains, en revanche, comprennent ce qu'est une situation sans issue et peuvent donc souvent élaborer une solution non optimale mais assez acceptable. Quand on cherche à savoir pourquoi une solution n'est pas valable, il est essentiel de prendre en compte la situation sans gagnant, car l'IA ne nous la présentera jamais.

Les IA que l'on utilise pour créer des solutions finiront par être à court d'idées, et elles deviendront alors globalement inutiles. La raison à cela est qu'une IA n'est pas créative. Les schémas sur lesquels elle opère existent déjà. Or, ces schémas ne répondent pas nécessairement aux besoins courants, ce qui revient à dire qu'il faut de nouveaux schémas. Les humains sont experts dans la création de nouveaux schémas pour la résolution des problèmes, et c'est en essayant encore que l'on peut créer de nouveaux schémas qu'une IA pourra utiliser pour nous rappeler quelque chose qui a fonctionné dans le passé. En résumé, le processus de résolution des problèmes est une boucle dans laquelle l'être humain est un maillon essentiel.

PARTIE 6
La Partie des Dix

DANS CETTE PARTIE…

» Trouver une activité dont l'IA ne vous dépossédera pas

» Découvrir ce que l'IA apporte à la société

» Comprendre pourquoi l'IA doit échouer dans certaines situations

Chapitre 18

Dix activités à l'abri de l'IA

DANS CE CHAPITRE

» **Interagir avec des humains**

» **Être créatif**

» **Utiliser son intuition**

Dans ce livre, nous avons beaucoup insisté sur la différence entre l'IA et l'être humain, et nous avons montré que les humains n'ont absolument aucune crainte à avoir. Certes, certains emplois vont disparaître, mais comme nous l'expliquons au Chapitre 17, l'utilisation de l'IA va générer tout un tas de nouveaux emplois, la plupart bien plus intéressants qu'un travail sur une chaîne de montage. Ces nouveaux emplois qui nous attendent concerneront des domaines de l'intelligence (voir Chapitre 1) auxquels une IA ne pourra tout simplement pas accéder. En réalité, compte tenu de l'incapacité de l'IA à maîtriser un certain nombre de domaines de la pensée humaine, beaucoup de gens garderont leur emploi actuel, et c'est le sujet de ce chapitre.



Votre métier actuel est à l'abri de l'IA s'il relève de certaines catégories dans lesquelles les relations humaines, la créativité et l'intuition jouent un rôle essentiel. Cependant, ce chapitre n'aborde que la partie visible de l'iceberg. L'alarmisme qui est propagé çà et là (voir

https://www.theinquirer.net/inquirer/news/3013919/elon-musk-spews-more-ai-fear-mongering-is-desperatebid-for-more-media-attention) fait que certains ont peur de voir leur emploi disparaître dans peu de temps. Par ailleurs, à cause de cet alarmisme, les gens n'exploitent pas tout le potentiel de l'IA, qui pourrait pourtant leur faciliter la vie. Le message général de ce chapitre est que vous n'avez rien à craindre et que l'IA, comme n'importe quel autre outil, sert à vous faciliter la vie et à vous permettre de vivre mieux.

Effectuer des interactions humaines

Les robots effectuent déjà des tâches comportant des interactions avec les humains, et ils en effectueront sans doute davantage dans l'avenir. Cependant, si l'on examine les applications habituelles de la robotique, on s'aperçoit que les robots exécutent essentiellement des tâches très rébarbatives : indiquer aux gens la direction à prendre, comme le ferait un employé d'un guichet d'information, ou rappeler aux personnes âgées que c'est le moment de prendre leurs médicaments, à l'image d'un réveille-matin, etc. Or, la plupart du temps, les interactions humaines ne sont pas aussi simples. Les sections qui suivent examinent les formes les plus interactives et les plus exigeantes d'interactions humaines : des activités qu'en aucun cas une IA ne serait capable d'exercer.

Enseigner aux enfants

Si vous pouviez passer un moment dans une classe d'école primaire, à observer la manière dont l'instituteur ou

l’institutrice s’occupe des enfants, vous seriez surpris. Pour schématiser, l’enseignant fait en sorte de conduire tous les enfants d’un point A à un point B avec un minimum de tracas, apparemment par la seule force de la volonté. Or, de façon inévitable, le niveau d’attention dont l’enfant a besoin n’est pas le même d’un enfant à l’autre. Tout ne se passe pas toujours bien, et l’enseignant est parfois amené à devoir résoudre plusieurs problèmes à la fois. Une IA, telle qu’elle peut exister aujourd’hui, serait dépassée par ce genre de situation, sachant que l’IA suppose une interaction humaine coopérative. Songeons une minute à ce que serait la réponse d’Alexa ou de Siri face à un enfant récalcitrant (ou face à un expérimentateur qui simulerait le même comportement). Cela ne pourrait pas fonctionner. Néanmoins, une IA peut assister l’enseignant dans les tâches suivantes :

- » classer des documents ;
- » utiliser un logiciel éducatif adaptif ;
- » améliorer les cours grâce à l’étude des modèles d’apprentissage ;
- » dispenser aux élèves des leçons particulières ;
- » montrer aux élèves comment trouver l’information ;
- » créer un environnement sûr pour l’apprentissage par essais et erreurs ;
- » guider les élèves dans leurs décisions concernant les cours et les activités extrascolaires à suivre en fonction de leurs aptitudes ;
- » aider les élèves à faire leurs devoirs.

Assurer des soins

Un robot peut servir à relever un patient, ce qui permet à l'infirmière d'éviter de s'esquinter le dos. En revanche, une IA ne peut pas décider quand, où ni comment il convient de relever le patient, car elle ne peut pas prendre en compte correctement toutes les informations requises, en particulier la communication non verbale du patient, ni comprendre sa psychologie, notamment une éventuelle tendance à dire des choses qui ne sont pas vraies (voir la section « Tenir compte des cinq types de données incorrectes » du Chapitre 2). L'IA peut poser des questions au patient, mais sans doute pas de la manière la plus appropriée pour obtenir des réponses exploitables. Un robot peut nettoyer les souillures, mais il ne saura sans doute pas le faire tout en préservant la dignité du patient et en lui donnant le sentiment qu'on prend soin de lui. En résumé, un robot est comme un bon marteau : très performant pour des tâches difficiles et grossières, mais pas particulièrement subtil ni attentionné.



De façon indubitable, le recours à l'IA sera un précieux apport à la profession médicale, mais ses applications restent très spécifiques et limitées. Le Chapitre 7 donne un bon aperçu de l'aide que l'IA peut apporter dans le domaine médical. Il s'agit d'activités qui ont rarement un rapport avec les interactions humaines. L'IA est plutôt affaire d'augmentation des capacités humaines et de collecte de données médicales.

Répondre à des besoins personnels

Peut-être considérez-vous votre IA comme la parfaite compagne. Il est vrai qu'elle ne répond jamais, qu'elle est toujours attentive et qu'elle ne vous quittera jamais pour quelqu'un d'autre. Vous pouvez lui confier vos pensées les plus intimes, et elle ne se moquera pas de vous. Une IA

comme Alexa ou Siri peut effectivement être une bonne compagne, comme dans le film *Her* (https://fr.wikipedia.org/wiki/Her). Le seul problème est qu'en réalité, une IA n'est pas du tout une bonne compagne. Tout ce qu'elle fait, c'est superposer une voix à un logiciel de navigation. L'anthropomorphisme ne fait pas de l'IA une véritable personne.

Le problème, quand une IA est censée répondre à des besoins personnels, c'est qu'elle ne comprend pas le concept de besoin personnel. Elle peut chercher une station de radio, trouver un article d'actualité, procéder à l'achat d'un produit, enregistrer un rendez-vous, vous avertir que le moment est venu de prendre vos médicaments, ou même, allumer et éteindre les lumières chez vous. Cependant, elle ne pourra pas vous dire si votre idée est bonne ou mauvaise, par exemple vous prévenir que ce que vous voulez faire risque de vous attirer de gros ennuis. Pour obtenir des conseils utiles dans des situations dans lesquelles il n'y a pas de règles définies, ou lorsque votre interlocuteur ne peut vous fournir une réponse ne serait-ce qu'approximative qu'en s'appuyant sur un vécu, il faut une personne humaine. C'est pour cela que le conseiller, le médecin, l'infirmière, ou même la dame que vous rencontrez occasionnellement dans un café et avec qui vous discutez, ne peuvent pas être remplacés par une machine. Certaines de ces personnes se font payer en monnaie sonnante et trébuchante, tandis que d'autres auront simplement besoin, à un moment donné, que vous les écoutiez à votre tour. Pour répondre à des besoins personnels qui ont vraiment un caractère personnel, les interactions humaines sont toujours indispensables.

Résoudre des problèmes de handicap

Quand les gens ont des besoins particuliers, il faut une touche humaine. Souvent, le besoin particulier se révèle être un cadeau particulier, mais seulement quand celui qui y répond le reconnaît comme tel. Parfois, une personne qui exprime un besoin particulier se porte très bien, et il faut de la créativité et de l'imagination pour découvrir le moyen de surmonter l'obstacle. Trouver le moyen de satisfaire un besoin particulier dans un monde qui n'accepte pas les besoins particuliers comme une chose normale est plus difficile encore. Ainsi, par exemple, la plupart des gens ne considèrent pas que le daltonisme puisse être un atout pour créer une œuvre d'art. Pourtant, quelqu'un a bien montré que c'était le cas (https://www.artsy.net/article/artsy-editorial-the-advantages-of-being-a-colorblind-artist).

Une IA peut être capable de venir en aide, de façon spécifique, à des personnes ayant des besoins particuliers. Un robot peut, par exemple, aider une personne à devenir mobile dans le cadre d'une ergothérapie ou d'une physiothérapie. Grâce à l'infinie patience du robot, la personne sera aidée de la même manière tous les jours. Cependant, seul un humain pourra s'apercevoir que la thérapie n'a pas d'effet et doit être changée.



Le handicap est un domaine dans lequel une IA, aussi bien programmée qu'elle puisse être, peut se révéler néfaste. Même si le patient réussit à effectuer certaines tâches, il se peut qu'il force trop, et seul un humain pourra s'en rendre compte grâce à divers messages non verbaux, ou bien grâce à son expérience ou à son intuition, deux qualités qu'une IA ne peut pas posséder en abondance car dans certaines situations, elle devrait extrapoler (étendre sa connaissance à une situation inconnue) plutôt qu'interpoler (utiliser sa connaissance entre deux limites connues) pour réussir. Pour résumer, il faut des humains

non seulement pour prendre soin du patient, conjointement avec l'IA, mais aussi pour vérifier que l'IA fonctionne comme prévu.

Créer de nouvelles choses

Comme on peut le constater à la lecture du Tableau 1.1, les robots ne peuvent rien créer. Il est essentiel de percevoir que la création consiste à élaborer de nouveaux schémas de pensée. Une bonne application d'apprentissage profond peut analyser les schémas de pensée existants, solliciter l'IA pour traduire ces schémas sous forme de nouvelles versions de ce qui existait déjà, et produire ce qui semblera être une pensée originale, mais sans qu'aucune créativité n'intervienne. Il ne s'agit jamais que de calculs mathématiques et de logique, pour analyser ce qui existe plutôt que définir ce qui pourrait exister. Compte tenu de ces limites de l'IA, les sections qui suivent décrivent la création de nouvelles choses, un domaine dans lequel ce sont toujours les humains qui excelleront.

Inventer

Quand on parle d'inventeurs, les gens pensent à des célébrités comme Thomas Edison, qui avait déposé 2332 brevets (dont 1093 aux États-Unis seulement) pour ses inventions (http://www.businessinsider.com/thomas-edisons-inventions-2014-2). Vous utilisez peut-être encore une de ses inventions, l'ampoule électrique, et un certain nombre d'entre elles, comme le phonographe, ont changé le monde. Tout le monde n'est pas Edison. Il y a eu aussi Bette Nesmith Graham (http://www.women-inventors.com/Bette-Nesmith-Graham.asp), qui a

inventé le correcteur liquide en 1956. Il fut un temps où son invention était utilisée dans tous les bureaux du monde pour corriger les fautes de frappe. Ces deux personnes ont fait quelque chose qu'une IA ne peut pas faire : créer un nouveau schéma de pensée sous forme d'une entité physique.



Certes, ces inventeurs ont trouvé leur inspiration dans d'autres sources, mais on leur doit réellement l'idée. Plus généralement, nous inventons continuellement. On peut trouver sur Internet des millions d'idées, qui proviennent toutes de gens qui ont simplement su voir quelque chose d'une manière différente. Une IA peut libérer les gens des tâches triviales et leur permettre ainsi de consacrer plus de temps à ce qu'ils peuvent faire le mieux : inventer toujours de nouvelles choses.

Faire de l'art

Le style et la présentation sont ce qui différencie un Picasso (https://www.pablopicasso.org/) d'un Monet (https://www.claudemonetgallery.org/). Nous pouvons voir la différence parce que nous distinguons des schémas dans la méthode de ces artistes, depuis le choix du support jusqu'à l'aspect de la peinture en passant par le style et les sujets traités. Une IA peut aussi cerner ces différences. En fait, sachant avec quelle précision une IA peut effectuer une analyse et combien de capteurs elle peut utiliser (dans la plupart des cas), elle peut probablement décrire le style d'une œuvre d'art mieux qu'un humain, et produire un résultat inédit dans le style même de l'artiste en question. Cependant, l'avantage qu'offre l'IA se limite à cela.



Une IA s'en tient à ce qu'elle connaît, tandis que les humains expérimentent. À l'adresse https://www.pinterest.com/aydeeyai/art-

made-with-non-traditional-materials-or-methods/, vous trouverez 59 exemples d'expérimentations réalisées uniquement avec des matériaux. Seul un humain aurait l'idée de créer une œuvre d'art avec du grillage (https://www.pinterest.fr/pin/47569339787387976/) ou des languettes de papier (https://www.pinterest.fr/pin/47569339789957402/). Des matériaux de toutes sortes ont déjà été utilisés pour créer des œuvres d'art qu'aucune IA ne pourra jamais reproduire.

Imaginer l'irréel

Les humains élargissent sans cesse l'ensemble du réel en rendant l'irréel possible. Il fut un temps où personne ne pensait qu'on volerait un jour dans des machines plus lourdes que l'air. Les expériences menées semblaient confirmer la théorie que le seul fait d'essayer de voler était une folie. Puis il y a eu les frères Wright (http://www.history.com/topics/inventions/wright-brothers). Leur vol à Kitty Hawk a changé le monde. Il importe cependant de se rendre compte que les frères Wright n'ont fait que rendre réel l'irréel qui occupait la pensée de beaucoup de gens (eux-mêmes y compris). Une IA ne pourrait pas produire un résultat irréel, et encore moins en faire une réalité. Seuls les humains peuvent le faire.

Prendre des décisions intuitives

L'*intuition* est une perception directe de la vérité, indépendante de tout raisonnement. C'est la vérité illogique, particulièrement difficile à analyser. Les humains sont doués pour cela, et les personnes les plus intuitives

disposent généralement d’un avantage significatif sur les autres. Or, l’IA, fondée sur la logique et les mathématiques, est dépourvue d’intuition. En conséquence, une IA doit généralement recenser toutes les solutions logiques possibles, et finalement conclure que le problème n’a pas de solution, même si un humain peut en trouver une avec une facilité relative. L’intuition et la perspicacité humaines jouent un rôle fondamental dans certaines activités, comme on va le voir dans les sections suivantes.

Enquêter sur les crimes

Souvent, dans les intrigues criminelles qu’on peut voir à la télévision, l’enquêteur parvient à résoudre toute l’affaire grâce à un détail. Dans la réalité, les choses se passent différemment. Les détectives se réfèrent à un savoir quantifiable, et il arrive que les criminels leur facilitent le travail. Les procédures et les méthodes établies jouent un rôle important dans la résolution des affaires criminelles. Il s’agit notamment de rechercher les preuves et de passer des heures rien qu’à les étudier. Or, il faut parfois que quelqu’un saute dans l’illogique pour que soudain, les pièces du puzzle s’assemblent.

Le travail d’un détective consiste à se confronter à tout un ensemble de problèmes. En fait, l’enquête ne porte pas nécessairement sur une activité illégale. Il arrive qu’un détective recherche simplement une personne qui a peut-être une bonne raison de vouloir rester introuvable. Surtout, il s’agit d’étudier des faits selon des méthodes auxquelles une IA ne penserait jamais parce qu’elles nécessitent le recours à l’intuition, une extension de l’intelligence qui n’existe pas dans l’IA. Il s’agit de *penser autrement*.

Contrôler des situations en temps réel

Une IA peut contrôler une situation en fondant les décisions futures sur des données antérieures. En d'autres termes, l'IA utilise des schémas pour formuler des prédictions. Dans la plupart des situations, cela ne pose pas de problème, et une IA peut réellement prédire avec un degré de précision élevé ce qui se produira dans un scénario particulier. Cependant, il arrive des situations dans lesquelles le schéma n'est pas valide et les données n'étayent pas la conclusion. Il se peut que les données manquent, ce qui se produit tout le temps. Dans ce genre de situation, l'intuition est le seul recours. En cas d'urgence, il n'est pas pertinent de compter uniquement sur une IA. Même si l'IA essaie la solution testée, quelqu'un peut penser autrement et trouver une autre idée.

Distinguer la réalité de la fiction

Une IA ne sera jamais intuitive. L'intuition va à l'encontre de toutes les règles actuellement utilisées pour créer une IA. En conséquence, des chercheurs ont décidé de créer Artificial Intuition (AN) (voir par exemple http://www.artificial-intuition.com/). À la lecture des documents à l'appui d'AN, il devient vite évident que les inventeurs comptent sur une sorte de magie (ils se bercent d'illusions) car la théorie ne correspond tout simplement pas à la mise en œuvre proposée.



AN pose des problèmes importants, en premier lieu le fait que tous les programmes, même ceux qui supportent l'IA, tournent sur des processeurs qui ne peuvent pas faire autre chose qu'exécuter les fonctions mathématiques et

logiques les plus simples. Qu'une IA puisse être aussi performante compte tenu du matériel actuellement utilisable n'est rien moins qu'étonnant.

Un second problème est que l'IA et tous les programmes informatiques exécutent les tâches au moyen d'opérations mathématiques. L'IA ne comprend rien. La section « Analyser l'argument de la chambre chinoise » du Chapitre 5 n'aborde qu'un des problèmes considérables que pose l'idée d'une capacité de compréhension de l'IA. En effet, l'intuition est illogique, si bien que les humains n'en comprennent même pas le fondement. Or, sans cette compréhension, il n'est pas possible de créer un système qui imite l'intuition, de quelque façon que ce soit.

Chapitre 19

Dix contributions importantes de l’IA à la société

DANS CE CHAPITRE

- » **Travailler avec les humains**
- » **Résoudre des problèmes industriels**
- » **Développer de nouvelles technologies**
- » **Effectuer des tâches dans l’espace**

Ce livre vous aide à comprendre l’histoire de l’IA et vous indique où elle est aujourd’hui, et où elle pourrait être demain. Cependant, une technologie n’est utile que dans la mesure où elle apporte une contribution substantielle à la société. Il faut aussi que cette contribution s’accompagne d’un fort incitatif financier, faute de quoi les investisseurs lui tourneront le dos. Même si le gouvernement peut contribuer à financer une technologie considérée comme utile à court terme à des fins militaires ou autres, la viabilité de la technologie à long terme repose sur le soutien des investisseurs. En conséquence, ce chapitre est consacré aux parties de l’IA qui sont utiles aujourd’hui, c’est-à-dire à celles qui représentent maintenant même une contribution notable à la société.



Selon certains, les promesses exagérées concernant les avantages de l’IA pourraient entraîner demain un nouvel hiver de l’IA.

En outre, les déclarations alarmistes de certaines personnalités influentes incitent les gens à remettre en question la valeur de l'IA (https://www.theinquirer.net/inquirer/news/3013919/elon-musk-spews-more-ai-fear-mongering-is-desperatebid-for-more-media-attention). D'autres estiment, au contraire, que le risque d'un hiver de l'IA est réduit (https://www.technologyreview.com/s/603062/ai-winter-isnt-coming/) et que cet alarmisme n'a pas lieu d'être.

Lorsqu'il s'agit d'évaluer une technologie, les discussions sont utiles, mais les paroles ne sont pas ce qui intéresse les investisseurs : ce qui les intéresse, ce sont les résultats. Ce chapitre parle des résultats et montre que l'IA est devenue une chose suffisamment intégrée dans la société pour que la survenue d'un autre hiver de l'IA soit peu probable. Bien sûr, il serait préférable que tout le battage sur l'IA prenne fin, pour que les gens puissent vraiment comprendre ce que l'IA peut faire pour eux.

S'intéresser à des interactions particulières avec les humains

C'est du public que dépendent les ventes des produits. C'est aussi le public qui décide des sujets dont on parlera le plus, qui crée le « buzz », et les ventes en sont la conséquence. Vous n'entendrez probablement pas parler à la radio des technologies abordées dans les sections suivantes, mais elles n'en ont pas moins un impact considérable sur le public. Dans le premier cas, une prothèse active de pied humain, les gens pourront effectivement marcher à l'aide de prothèses, presque aussi facilement qu'on marche avec ses vrais pieds. Même si ce produit s'adresse à un groupe relativement

restreint de personnes, les effets peuvent être largement connus. Le deuxième et le troisième cas sont susceptibles d'affecter des millions de gens, peut-être même des milliards. Ce sont des offres banales, mais ce qui est banal est souvent ce que les gens attendent, et c'est, là encore, ce qui fait vendre. Dans les trois cas, les technologies ne fonctionneront pas sans IA, ce qui veut dire que si la recherche-développement et les ventes s'arrêtent, les utilisateurs de ces technologies ne seront sans doute pas très contents.

Concevoir une prothèse active de pied humain

Les prothèses, c'est beaucoup d'argent. Elles coûtent une fortune à fabriquer, or ce sont des objets absolument nécessaires pour quiconque a perdu un membre et veut conserver une qualité de vie décente. Les prothèses sont souvent passives, ce qui signifie qu'elles ne réagissent pas et n'adaptent pas automatiquement leur fonctionnalité aux besoins de la personne, mais tout cela est en train de changer depuis quelques années. Des scientifiques comme Hugh Herr (https://www.smithsonianmag.com/innovation/future-robotic-legs-180953040/) ont créé des prothèses actives, c'est-à-dire capables de simuler les mouvements des membres de la personne et de s'adapter automatiquement à la personne qui les utilise. Hugh Herr a fait les gros titres des journaux, mais les technologies actives sont aujourd'hui présentes dans toutes sortes de prothèses, notamment les prothèses de genou, de bras et de main.



Vous vous demandez peut-être quelle valeur potentielle représente l'utilisation d'une prothèse active, par rapport à une prothèse passive. Les fabricants se sont déjà lancés

dans ce genre de recherche (voir les résultats d'un rapport : https://www.rand.org/pubs/research_reports/RR2096.html). Il s'avère que les prothèses à base de microprocesseur assistées d'une IA assurant une interaction correcte avec l'utilisateur représentent un enjeu considérable. Non seulement leurs utilisateurs vivent plus longtemps, mais ces prothèses réduisent aussi les frais médicaux directs et indirects. Par exemple, une personne qui utilise une prothèse à technologie active risque moins de tomber. Même si le coût initial est plus élevé, les coûts à moyen et long terme sont bien plus réduits.

Assurer un contrôle permanent

Le Chapitre 7 présente un ensemble de dispositifs de contrôle utilisés en médecine pour que les patients prennent bien leurs médicaments en temps et en heure, et que les doses soient respectées. Par ailleurs, ce type de suivi médical permet aux patients de recevoir les soins plus rapidement en cas d'incident majeur et il permet même de prédire chez le patient un incident majeur, par exemple une crise cardiaque. La plupart de ces dispositifs, et plus particulièrement ceux qui sont prédictifs par nature, fonctionnent grâce à une IA. Cependant, la question reste posée de savoir si ces systèmes sont un incitatif financier pour leurs créateurs comme pour leurs utilisateurs.

Les études sont peu nombreuses, mais les résultats présentés sur la page https://academic.oup.com/europace/article-abstract/19/9/1493/3605206 montrent que la surveillance à distance des patients atteints de pathologies cardiaques permet de réduire considérablement les coûts des traitements (et elle permet aussi aux patients de vivre mieux et plus longtemps). D'après le *Financial Times*

(https://www.ft.com/content/837f904e-9fd4-11e4-9a74-00144feab7de), le recours à un système de surveillance à distance, même pour les individus en bonne santé, a une incidence importante sur les coûts médicaux (l'article est réservé aux abonnés). L'impact des économies réalisées est tel que le contrôle à distance est en train de révolutionner la médecine.

Administrer les médicaments

Quand les patients oublient de prendre leurs médicaments, cela revient très cher à la société. Aux États-Unis seulement, d'après un article publié par CNBC.com (*https://www.cnbc.com/2016/08/03/patients-skipping-meds-cost-290-billion-per-year-can-smart-pills-help.html*), cela représente un coût de 290 milliards de dollars par an. En combinant des technologies comme Near Field Communication (NFC) (https://www.nfcworld.com/2015/11/18/339766/nxp-launchesnfcblister-packs-and-pill-bottles-for-medication-tracking/) avec des applications mobiles à base d'IA, on peut vérifier la manière dont les patients prennent leurs médicaments et le moment où ils les prennent. Par ailleurs, l'IA peut leur signaler le moment où ils doivent les prendre, et leur rappeler quel médicament ils doivent prendre et à quelle dose. En associant ces applications à un dispositif de suivi, même les patients qui ont besoin d'une surveillance particulière absorberont la dose adéquate de chaque médicament (https://clinicaltrials.gov/ct2/show/NCT02243670).

Développer des solutions industrielles

Le public effectue des petits achats en quantités énormes. Or, le pouvoir d'achat d'un individu apparaît bien limité par rapport à ce qu'une organisation est capable de dépenser. La différence est dans la quantité. Les investisseurs s'intéressent aux deux types de ventes, car dans les deux cas, l'enjeu financier est considérable. Les solutions industrielles ont une incidence sur les organisations. Elles sont généralement onéreuses, mais l'industrie y recourt pour accroître la productivité, l'efficience, et avant tout, les revenus. Tout est affaire de profit. Les sections qui suivent étudient l'incidence de l'IA sur les profits des organisations.

Utiliser l'IA pour l'impression 3D

L'impression 3D a d'abord été une technologie ludique produisant des résultats intéressants mais pas particulièrement exploitables. Cependant, c'était avant que la NASA utilise l'impression 3D dans la station spatiale internationale (SSI) pour fabriquer des outils (https://www.nasa.gov/content/international-space-station-s-3-d-printer). On aurait pu penser que la SSI, en quittant la Terre, emporterait tous les outils nécessaires. Malheureusement, les outils se perdent et se brisent. Par ailleurs, la SSI ne dispose pas d'un espace suffisant pour stocker tous les outils dont on peut avoir besoin sur place. L'impression 3D permet aussi de fabriquer des pièces de rechange, or la SSI ne peut certainement pas en emporter un assortiment complet. En microgravité, les imprimantes 3D fonctionnent de la même

manière que sur terre (https://www.nasa.gov/mission_pages/station/research/experiments/1115.html), par conséquent, cette technologie peut être utilisée par les scientifiques de la même manière sur terre et dans l'espace.

Dans le même temps, l'industrie utilise l'impression 3D pour répondre à toutes sortes d'exigences. Quand la machine est complétée par une IA, elle produit un résultat, elle analyse ce résultat, et elle apprend de ses erreurs (*https://www.digitaltrends.com/cool-tech/ai-build-wants-to-change-the-way-we-build-the-future/*). Par conséquent, l'industrie pourra un jour ou l'autre fabriquer des robots qui corrigeront leurs propres erreurs, du moins dans une certaine mesure, si bien qu'il y aura moins d'erreurs et plus de profits. L'IA permet aussi de réduire le risque associé à l'impression 3D, avec des produits comme Business Case (*https://www.sculpteo.com/blog/2017/08/10/the-artificial-intelligence-for-your-3d-printing-projects-business-case/*).

Faire progresser les technologies robotiques

Ce livre contient beaucoup d'informations sur l'utilisation des robots, aussi bien chez les particuliers qu'en médecine et dans l'industrie. Nous vous parlons aussi des robots utilisés dans les véhicules, dans l'espace et sous l'eau. Vous vous en doutez, les robots jouent un rôle significatif dans la demande d'IA. Ils sont de plus en plus fiables et abordables, leur technologie est de mieux en mieux maîtrisée, leur présence devient visible et ils ont fait leurs preuves, ce qui explique que tant d'organisations investissent dans des robots toujours plus sophistiqués.

Aujourd'hui, les robots sont utilisés dans un certain nombre d'activités traditionnelles, une chose que le public ne sait pas nécessairement. L'industrie pétrolière, par exemple, est aujourd'hui fortement tributaire des robots pour la recherche de nouveaux gisements, pour la maintenance des installations et pour l'inspection des conduites. Dans certains cas, les robots effectuent aussi des réparations en des lieux difficilement accessibles pour les humains, notamment à l'intérieur des oléoducs (http://insights.globalspec.com/article/2772/the-growingroleof-artificial-intelligence-in-oil-and-gas). D'après Oil & Gas Monitor, l'IA rend possible l'interpolation entre les modèles d'exploitation minière, réduit les coûts de forage et effectue des simulations faisant apparaître les problèmes de forage potentiels (http://www.oilgasmonitor.com/artificial-intelligence-upstream-oil-gas/). L'utilisation de l'IA permet aux ingénieurs de réduire le risque de façon générale, ce qui signifie que l'incidence des hydrocarbures sur l'environnement est susceptible de se réduire grâce à une diminution des déversements.



D'après Engineering 360, la baisse des cours des hydrocarbures est ce qui a conduit l'industrie pétrolière à adopter l'IA (http://insights.globalspec.com/article/2772/the-growing-role-of-artificial-intelligence-in-oil-and-gas). L'industrie pétrolière présente une forte aversion au risque, aussi recourt-elle à l'IA et cette expérience a valeur de test pour savoir de quelle manière d'autres secteurs d'activité l'adopteront également. En passant en revue les articles publiés sur l'industrie pétrolière, on se rend compte que celle-ci a attendu d'assister à des succès dans les secteurs de la médecine, de la finance et de l'industrie manufacturière avant d'investir à son tour dans cette

technologie. À mesure que l'IA fera ses preuves dans d'autres secteurs, on peut s'attendre à ce que son adoption connaisse un nouveau regain.



Ce livre traite de toutes sortes de solutions robotiques, mobiles ou non. La Quatrième partie traite des robots en général, des robots volants (car c'est précisément ce que sont les drones, quand on y réfléchit), et des voitures sans conducteur. De façon générale, les robots sont générateurs de profit quand ils effectuent certaines tâches spécifiques, comme nettoyer le sol de votre appartement (le Roomba) ou assembler les éléments d'un véhicule. De même, les drones sont aujourd'hui rémunérateurs pour les entreprises qui travaillent pour la défense, et ils deviendront rentables tôt ou tard dans un certain nombre d'utilisations civiles. Selon nombre de prédictions, la voiture sans conducteur va devenir non seulement rentable mais aussi très populaire (*https://www.forbes.com/sites/oliviergarret/2017/03/03/10-million-self-driving-cars-will-hit-the-road-by-2020-heres-how-to-profit/#7af725d87e50*).

Créer de nouveaux environnements technologiques

Le public est généralement friand de nouveautés à acheter, aussi les entreprises doivent-elles proposer de nouveaux produits sur le marché. L'IA permet de repérer des tendances dans toutes sortes de domaines. C'est souvent ainsi que l'on peut découvrir la présence d'une nouveauté, par exemple un nouvel élément ou un nouveau processus de fabrication. Dans le domaine du développement de produits, le rôle de l'IA est d'aider à la découverte du nouveau produit (par opposition aux activités centrées sur la vente d'un produit existant). En réduisant le délai nécessaire pour trouver un nouveau

produit à commercialiser, l’IA permet aux entreprises d’accroître leurs profits et réduit le coût de la recherche afférente. La section suivante aborde ces questions plus en détail.

Développer de nouvelles ressources rares

Comme le montre ce livre, une IA est particulièrement performante pour identifier des motifs ou des tendances, et les tendances peuvent indiquer toutes sortes de choses, notamment de nouveaux éléments (la section « Trouver de nouveaux éléments » du Chapitre 16 est consacrée à cet aspect de l’IA). De nouveaux éléments, cela signifie de nouveaux produits, et donc de nouvelles ventes. Une entreprise qui peut proposer un nouveau matériel dispose d’un avantage significatif par rapport à la concurrence. L’article de la page https://virulentwordofmouse.wordpress.com/2010/11/30/an-economic-perspective-on-revolutionary-us-inventions/ nous éclaire sur l’impact économique de certaines inventions parmi les plus intéressantes. Souvent, ces inventions mettent en jeu un nouveau procédé ou un nouveau matériau que l’IA a permis assez facilement de trouver.

Voir ce qui ne peut pas être vu

La vision humaine ne perçoit pas l’ensemble du spectre lumineux existant dans la nature. Même avec une augmentation « via une loupe ou un zoom par exemple », nous avons des difficultés à envisager ce qui est à très petite échelle ou à très grande échelle. Certains biais nous empêchent de voir l’inattendu. Un modèle aléatoire possède parfois une structure que nous ne pouvons pas

déceler. Une IA peut voir ce que les humains ne peuvent pas voir, et elle peut agir en conséquence. Elle peut, par exemple, repérer une contrainte exercée sur du métal et intervenir pour éviter que la pièce métallique en question se déforme. Les économies peuvent être monumentales lorsqu'il s'agit, par exemple, de guides d'ondes comme ceux utilisés pour les transmissions radiophoniques (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4481976/).

Travailler avec l'IA dans l'espace

Le Chapitre 16 vous fait découvrir le rôle que l'IA est susceptible de jouer dans l'espace. Si des plans pour la réalisation de ces tâches sont à l'étude, la plupart sont soutenus par les gouvernements, ce qui signifie qu'il s'agit d'opportunités qui ne déboucheront pas nécessairement sur du profit. Au Chapitre 16, il est aussi question de projets de recherche dans le secteur privé. Les entreprises cherchent à réaliser un profit, mais pas nécessairement dans l'immédiat. Les sections qui suivent offrent un autre regard sur l'espace et montrent ce qui se produit actuellement. L'IA va permettre aux entreprises de gagner de l'argent en travaillant dans l'espace, et cela les incite à continuer d'investir à la fois dans l'IA et dans des projets spatiaux.

Livrer des marchandises aux stations spatiales

Le plus grand succès commercial de l'IA dans l'espace, à ce jour, est probablement le réapprovisionnement de la SSI par des compagnies comme SpaceX et Orbital ATK (https://www.nasa.gov/mission_pages/station

/structure/launch/overview.html). Ces compagnies gagnent de l'argent à chaque expédition, bien évidemment, mais la NASA en profite également. Ce sont les États-Unis dans leur ensemble, qui tirent profit de cette entreprise :

» une réduction des coûts de livraison des marchandises, en évitant d'utiliser des engins spatiaux d'autres pays pour réapprovisionner la SSI ;

» une utilisation accrue des installations américaines, comme le Kennedy Space Center, le coût de ces installations étant amorti sur une longue période ;

» de nouveaux centres de lancement pour les futurs vols dans l'espace ;

» une plus grande capacité de charge utile pour les satellites et autres engins.

SpaceX et Orbital ATK ont des liens avec beaucoup d'autres entreprises. Par conséquent, même s'il semble que deux compagnies seulement profitent de ce dispositif, bien d'autres en profitent également en tant que partenaires affiliés. C'est l'utilisation de l'IA qui rend tout cela possible, et cela se produit en ce moment même. Des entreprises gagnent d'ores et déjà de l'argent avec l'espace, sans attendre demain, contrairement à ce que les médias pourraient nous faire croire. Le fait que les profits soient dégagés par ce qui est essentiellement un vulgaire service de livraison ne fait aucune différence.

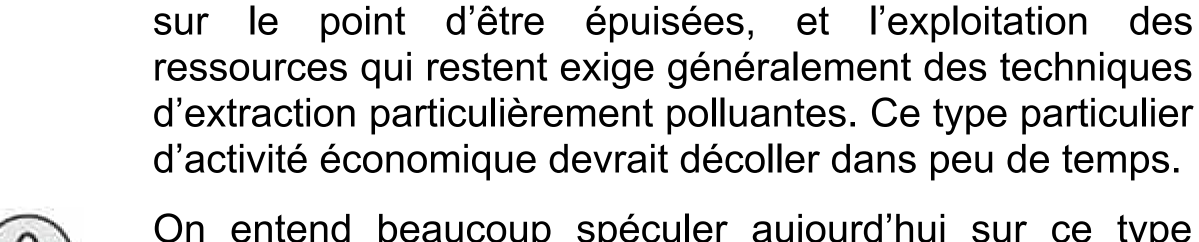


Les livraisons dans l'espace sont, pour l'essentiel, quelque chose de nouveau. De nombreuses entreprises ayant leurs activités sur Internet ont connu plusieurs déficits annuels successifs avant de devenir rentables. SpaceX, au moins, semble être sur le point de gagner de l'argent après quelques années de pertes initiales (https://www.fool.com/investing/2017/02/05/how-profitable-is-spacex-really.aspx). Les

entreprises spatiales auront besoin de temps pour obtenir des résultats financiers comparables à ceux qu'affichent aujourd'hui les entreprises qui exercent les mêmes activités au sol.

Exploiter des ressources extraterrestres

Des compagnies comme Planetary Resources (https://www.planetaryresources.com/) sont sur le point de commencer à exploiter du minerai sur des astéroïdes et autres corps planétaires. De façon certaine, c'est là que se trouvent les possibilités de gagner d'énormes sommes d'argent (http://theweek.com/articles/462830/how-asteroid-mining-could-add-trillions-world-economy). Nous incluons cette section dans ce chapitre car les ressources minières terrestres sont littéralement sur le point d'être épuisées, et l'exploitation des ressources qui restent exige généralement des techniques d'extraction particulièrement polluantes. Ce type particulier d'activité économique devrait décoller dans peu de temps.





On entend beaucoup spéculer aujourd'hui sur ce type particulier d'activité, entre autres à propos de l'exploitation de (16) Psyché (https://www.usatoday.com/story/tech/nation-now/2017/01/18/nasa-planning-mission-asteroid-worth-10000-quadrillion/96709250/). Quoi qu'il en soit, il serait nécessaire de mettre sur pied un formidable programme de recyclage, qui semble irréaliste, ou bien trouver des ressources ailleurs, très probablement dans l'espace. Actuellement, ceux qui gagnent de l'argent sur ce projet particulier sont ceux qui fournissent les outils permettant de déterminer la meilleure manière de le réaliser, et ces outils sont bien souvent à base d'IA.

Explorer d’autres planètes

Il semble probable que l’homme finisse par explorer et même par coloniser d’autres planètes, à commencer par Mars. Déjà 78000 personnes se sont inscrites pour effectuer ce voyage (voir http://newsfeed.time.com/2013/05/09/78000-people-apply-for-one-way-trip-to-mars/). Une fois que nous aurons atteint d’autres mondes, y compris la Lune, beaucoup de gens pensent que la seule façon de gagner de l’argent consistera à vendre de la propriété intellectuelle, ou peut-être de créer des matériaux particuliers à ces lieux de l’espace (https://www.forbes.com/sites/quora/2016/09/26/is-there-a-fortune-to-be-made-on-mars/#68d630ab6e28).



Malheureusement, si certaines personnes gagnent déjà de l’argent avec ce projet, nous ne pouvons pas nous attendre à ce que nos efforts soient rentables avant un bon moment. Toujours est-il que certaines compagnies font aujourd’hui du profit en produisant les divers outils nécessaires à la conception d’une telle expédition. C’est la recherche qui finance l’économie.

Chapitre 20

Dix exemples d'échecs de l'IA

DANS CE CHAPITRE

» **Concevoir le monde**

» **Développer de nouvelles idées**

» **Comprendre la condition humaine**

Tout ouvrage qui ambitionne de traiter de façon complète de l'IA doit parler des cas dans lesquels l'IA n'a pas répondu aux attentes. Ce sujet a été en partie abordé dans d'autres chapitres, dans le cadre de l'historique des hivers de l'IA. Cependant, ces considérations ne suffisent peut-être pas à faire comprendre pourquoi l'IA n'a pas toujours répondu aux attentes exprimées par ses promoteurs enthousiastes. En effet, l'IA n'a pas satisfait à certains besoins spécifiques ni à certaines exigences de base. Ce chapitre parle des échecs qui empêcheront l'IA d'exceller et d'accomplir les tâches qu'elle devrait accomplir pour que se réalisent pleinement des succès comme ceux mentionnés dans d'autres chapitres. L'IA est actuellement une technologie en pleine évolution, qui n'est que partiellement couronnée de succès dans le meilleur des cas.



Un des problèmes essentiels qui se posent aujourd'hui avec l'IA est que les gens continuent à faire de l'anthropomorphisme à son sujet et à en faire ce qu'elle n'est pas. Une IA accepte en entrée des données

nettoyées, elle les analyse, identifie des tendances, et produit un résultat requis en sortie. Comme l'expliquent les sections de ce chapitre, une IA ne comprend rien, ne peut rien créer ni découvrir de nouveau et elle n'a aucune connaissance intrapersonnelle, si bien qu'elle ne peut pas faire preuve de compréhension vis-à-vis de quelqu'un à propos de quelque chose. L'information la plus essentielle à retenir de ce chapitre, c'est que le comportement d'une IA est conçu par un programmeur humain et que ce que l'on prend souvent pour de l'intelligence n'est que l'association d'une programmation intelligente et de vastes quantités de données analysées d'une façon spécifique. Pour un autre point de vue sur ces questions, entre autres, consultez l'article intitulé « Asking the Right Questions About AI » sur la page https://medium.com/@yonatanzunger/asking-the-right-questions-aboutai-7ed2d9820c48.

Cependant, plus important encore est le fait que certains affirment que l'IA finira par prendre le contrôle du monde, sans comprendre que c'est impossible en l'état actuel des technologies. Une IA ne peut pas soudainement accéder à la conscience d'elle-même, car elle ne dispose d'aucun moyen de ressentir des émotions, ce qui serait une condition pour y accéder. Aujourd'hui, comme l'indique le Tableau 1.1 du Chapitre 1, plusieurs des sept types fondamentaux d'intelligence requis pour avoir la conscience de soi font défaut aux IA. Par ailleurs, être simplement doté de ces sept niveaux d'intelligence n'y suffirait pas. Nous avons aussi en nous ce que l'on appelle une étincelle de vie, une chose que les scientifiques ne comprennent pas. Or, sans la compréhension de ce qu'est cette étincelle, la science ne peut pas la recréer dans le cadre d'une IA.

Comprendre

La faculté de compréhension est innée chez les humains, mais les IA en sont totalement dépourvues. Pour un humain, une pomme n'est pas seulement une série de propriétés associées à l'image d'un objet. Nous comprenons ce qu'est une pomme grâce à l'utilisation de nos sens : la vue, le goût et le toucher. Nous comprenons que la pomme est comestible et que sa consommation nous fournit certains éléments nutritifs. Nous ressentons des choses à propos des pommes : nous pouvons les aimer, elles peuvent être notre fruit préféré. L'IA perçoit un objet auquel des propriétés sont associées : ce sont des valeurs qu'elle manipule mais qu'elle ne comprend pas. Les sections qui suivent expliquent pourquoi cette incapacité de comprendre empêche l'IA, globalement, de satisfaire à certaines attentes.

Interpréter plutôt qu'analyser

Comme on l'a vu à maintes reprises dans ce livre, une IA utilise des algorithmes pour manipuler les données entrantes et produire un résultat. L'accent est mis sur l'analyse des données. Or, c'est un humain qui contrôle l'orientation donnée à cette analyse et qui doit ensuite interpréter les résultats. Ainsi, par exemple, une IA peut analyser une radiographie indiquant une tumeur potentiellement cancéreuse. Le résultat peut mettre en évidence la zone particulière du cliché dans laquelle se trouve la tumeur, afin que le médecin la voie bien. Autrement, il ne pourrait peut-être pas la déceler. Il est donc indéniable que l'IA rend là un service important. Cependant, il faut encore que ce soit un médecin qui examine le résultat et détermine si la radio montre vraiment un cancer. Comme l'expliquent plusieurs sections de ce livre, en particulier celle consacrée aux voitures sans conducteur au Chapitre 14, une IA peut facilement être induite en erreur lorsqu'un petit artéfact apparaît au

mauvais endroit. Par conséquent, même si l'IA est extrêmement utile pour permettre au médecin de voir ce qui n'est pas visible directement à l'œil humain, elle n'est pas suffisamment fiable pour qu'une décision d'aucune sorte puisse lui être confiée.

L'interprétation suppose aussi la capacité de voir au-delà des données. Il s'agit de la capacité non pas de créer de nouvelles données, mais de comprendre que les données peuvent indiquer autre chose que ce qui est apparent. Ainsi, nous pouvons souvent nous apercevoir que certaines données sont fausses ou falsifiées, même en l'absence de preuve. Une IA accepte les données et les considère comme réelles et vraies, même lorsqu'un être humain ne peut ignorer qu'elles ne sont ni réelles ni vraies. Il est actuellement impossible de formaliser avec précision la manière dont l'être humain y parvient, sachant que c'est une chose que nous ne comprenons pas nous-mêmes.

Aller au-delà des chiffres

Malgré l'apparence éventuelle du contraire, une AI travaille uniquement avec des nombres. Elle ne peut pas comprendre des mots, par exemple. Quand vous lui parlez, l'IA ne fait que rechercher des correspondances après avoir transcrit vos paroles sous une forme numérique. La substance de votre message lui échappe. Même si elle était capable de comprendre les mots, elle ne saisirait pas votre message, car les mots disparaissent à l'issue du processus de transformation en unités lexicales. L'IA étant incapable de comprendre une chose aussi fondamentale que les mots, il manquera toujours à la traduction d'un texte d'une langue à une autre par une IA ce quelque chose qui permet de traduire un sentiment, aussi bien que les mots mêmes. Les paroles expriment des sentiments, ce qu'une IA ne peut pas faire.

Il se produit le même processus de conversion avec chacun de nos sens. Un ordinateur traduit la vue, le son, l'odeur, le goût et le toucher sous forme de représentations numériques, puis réalise une correspondance entre des schémas pour créer un jeu de données qui simule l'expérience réelle. Pour compliquer encore les choses, les humains ne perçoivent souvent pas tous les choses de la même manière. Par exemple, chacun a sa propre perception des couleurs (https://www.livescience.com/21275-color-red-blue-scientists.html). Du point de vue d'une IA, tout ordinateur voit la couleur de la même façon exactement, ce qui signifie qu'une IA ne peut pas avoir une expérience unique des couleurs. En outre, en raison du processus de conversion, l'IA n'a en réalité aucune expérience de la couleur.

PRENDRE EN COMPTE LE COMPORTEMENT HUMAIN

L'absence de compréhension du comportement humain est un thème récurrent dans ce chapitre. Même la compréhension d'un comportement ne suffit pas pour le reproduire ou le simuler. Pour qu'une IA puisse le prendre en compte, il faut qu'il soit mis sous une forme mathématique. Sachant que de très nombreux comportements humains ne sont pas du tout compris, il y a peu de chances que quelqu'un puisse créer un modèle mathématique formel pour cela avant longtemps. Or, sans de tels modèles, une IA ne peut pas penser comme un être humain ni atteindre quoi que ce soit qui ressemble au sentiment ou à la conscience de soi.

Considérer les conséquences

Une IA peut analyser des données, mais elle ne peut pas produire un jugement moral ou éthique. Si vous demandez à une IA de faire un choix, elle choisira toujours l'option présentant la plus forte probabilité de succès, à moins de prévoir une fonction de randomisation. L'IA procédera à ce choix sans se soucier du résultat. L'encadré « **Les voitures sans conducteur et le dilemme du tramway** », au Chapitre 14, présente ce problème très clairement. Confrontée au choix entre la mort des passagers d'une voiture et la mort des piétons, lorsque cette alternative est inévitable, l'IA a besoin que l'être humain lui donne des instructions pour pouvoir prendre la décision. L'IA n'est pas capable de tenir compte des conséquences, aussi n'est-elle pas apte à participer au processus de décision.



Dans diverses situations, mal évaluer la capacité d'une IA à exécuter une tâche n'est qu'un inconvénient mineur. Dans certains cas, il faudra exécuter la tâche une deuxième ou une troisième fois manuellement, l'IA n'étant pas compétente. Cependant, en ce qui concerne les conséquences, si vous vous en remettez à une IA pour exécuter une tâche pour laquelle elle n'est pas compétente, vous risquez de vous retrouver confronté à des problèmes d'ordre juridique, outre les problèmes moraux et éthiques. Par exemple, si vous faites circuler une voiture sans conducteur dans un périmètre qui n'est pas prévu pour cela, ce sera sans doute illicite et vous serez confronté à des problèmes juridiques, sans compter les dégâts et les frais médicaux éventuels. En un mot, il convient de connaître les exigences légales avant de confier à une IA une tâche susceptible d'avoir des conséquences.

Découvrir

Une IA peut interpoler des connaissances existantes, mais elle ne peut pas extrapoler pour générer une

connaissance nouvelle. En principe, quand une IA se trouve confrontée à une situation nouvelle, elle essaye de résoudre le problème en considérant les conditions courantes comme connues, au lieu de considérer que c'est quelque chose de nouveau. Une IA ne dispose d'aucune méthode pour créer, ni pour voir une situation comme unique. Ce sont des expressions humaines qui nous permettent de découvrir de nouvelles choses, de travailler avec ces nouvelles choses, de concevoir des méthodes pour interagir avec ces nouvelles choses, et de créer de nouvelles méthodes afin d'utiliser ces nouvelles choses pour exécuter de nouvelles tâches ou augmenter des tâches existantes. Les sections qui suivent montrent que l'IA, parce qu'elle est incapable de faire des découvertes, ne peut pas faire ce que les humains attendent d'elle.

Faire de nouvelles données avec les anciennes

L'*extrapolation* des données est une des tâches les plus courantes que nous effectuons : par exemple, connaissant A, que vaut B ? Nous utilisons nos connaissances existantes pour créer une connaissance d'un type différent. À partir de ce que nous savons, nous pouvons nous aventurer hors du domaine de notre connaissance initiale avec une probabilité élevée de réussite. Nous faisons cela si souvent que cela devient pour nous une seconde nature. Le processus est extrêmement intuitif. Même les enfants peuvent se livrer à ces prédictions avec un taux de réussite élevé.



Une IA ne pourra jamais faire mieux qu'*interpoler* des données. Connaissant A et B, est-ce que C se trouve quelque part entre les deux ? C'est là un exemple de problème d'interpolation. Compte tenu de cette capacité

d’interpolation des données, une IA peut étendre un schéma, mais elle ne peut pas créer de nouvelles données. Cependant, il arrive que des développeurs laissent le public croire que les données sont nouvelles, grâce à des techniques de programmation subtiles. La présence de C paraît nouvelle, mais en réalité elle ne l’est pas. En l’absence de nouvelles données, les conditions peuvent être réunies pour que l’IA semble résoudre un problème, mais c’est une illusion. Le problème nécessite une solution nouvelle, et non pas l’interpolation de solutions existantes.

Voir au-delà des schémas

Actuellement, une IA peut identifier dans les données des schémas que les humains ne voient pas. La capacité de percevoir ces schémas est ce qui rend l’IA si intéressante. La manipulation et l’analyse des données demandent du temps et ce sont des activités complexes et répétitives, mais une IA peut s’en charger sans problème. Or, les schémas décelés dans les données sont un résultat, mais pas nécessairement une solution. Nous mobilisons nos cinq sens, notre empathie, notre créativité et notre intuition pour voir au-delà des schémas et envisager une solution potentielle hors du champ de ce que les données pourraient faire croire. Le Chapitre 18 traite plus en détail de cet aspect de la condition humaine.



Voici un moyen simple de comprendre la capacité humaine de voir au-delà des schémas. Quand on observe un ciel nuageux, on peut voir des dessins formés par les nuages, tandis qu’une IA ne verra que des nuages et rien d’autre. En outre, deux personnes qui observeront les mêmes nuages pourront y voir des choses différentes, car leur vision est créative : l’une verra un mouton là où une autre verra une fontaine. Il en est de même pour un ciel étoilé et d’autres sortes de schémas. L’IA présente le

schéma en guise de résultat, mais elle n'a aucune compréhension de ce schéma et elle n'a pas la créativité qui lui permettrait d'en faire quelque chose, à part en révéler l'existence.

Mettre en œuvre de nouveaux sens

En devenant plus instruits, nous sommes aussi devenus conscients de certaines variations des sens de l'être humain qui ne peuvent pas bien être traduites au niveau d'une IA car il n'est pas vraiment possible aujourd'hui de reproduire ces sens en informatique. Ainsi, par exemple, la capacité d'utiliser plusieurs sens pour gérer une entrée unique (la synesthésie), voir (*https://www.mnn.com/health/fitness-well-being/stories/what-is-synesthesia-and-whats-it-like-to-have-it*) dépasse ce que l'on peut envisager pour une IA.

Décrire comme il convient la synesthésie est au-delà des possibilités de la plupart des gens. Or, avant de pouvoir créer une IA qui imite une partie des effets remarquables de la synesthésie, il faut tout d'abord la décrire entièrement, puis créer des capteurs qui traduiront l'expérience sous forme de nombres qu'une IA pourra analyser. Cependant, même dans ce cas, l'IA ne percevra que les effets de la synesthésie, et non l'impact émotionnel. Par conséquent, une IA ne pourra jamais faire pleinement l'expérience d'une synesthésie ni la comprendre (la section « Modifier le spectre des données » du Chapitre 8 explique de quelle manière une IA peut augmenter la perception humaine grâce à une expérience de synesthésie). Assez curieusement, certaines études montrent qu'il est possible d'habituer des adultes à avoir des expériences de synesthésie, si bien

que le besoin d'une IA est incertain (https://www.nature.com/articles/srep07089).



S'il est généralement connu que l'être humain est doté de cinq sens, diverses sources affirment aujourd'hui qu'en réalité, il en posséderait bien davantage (*http://www.todayifoundout.com/index.php/2010/07/humans-have-a-lot-more-than-five-senses/*). Ces sens supplémentaires ne sont pas bien cernés et il est difficile de prouver leur existence. C'est le cas, notamment, de la *magnétoréception* (la capacité de détecter des champs magnétiques, entre autres le champ magnétique terrestre). Ce sens, similaire à celui des oiseaux mais à un moindre degré, nous permettrait de nous orienter. Comme nous ne disposons d'aucune méthode pour quantifier ce sens, le reproduire dans le cadre d'une IA est impossible.

S'identifier à quelqu'un

Les ordinateurs n'éprouvent rien. Ce n'est pas nécessairement une mauvaise chose, mais dans ce chapitre, cet aspect est effectivement négatif. Sans la capacité de ressentir quelque chose, l'ordinateur ne peut pas voir les choses du point de vue d'un être humain. Il ne comprend pas ce qu'est le bonheur, ni la tristesse, et il ne peut donc pas réagir à ces émotions, à moins qu'un programme crée une méthode lui permettant d'analyser les expressions faciales et d'autres indicateurs, et à agir en conséquence. Même dans ce cas, la réaction en question ne serait qu'une réponse préprogrammée et comportant un risque d'erreur. Songez à toutes les décisions que vous prenez en fonction d'un besoin émotionnel plutôt que d'un fait objectif. Les sections qui suivent montrent comment l'absence d'empathie de la part d'une IA fait que dans de nombreux cas, elle ne peut pas interagir avec les humains de façon appropriée.

Se mettre dans la peau de quelqu’un

Se mettre dans la peau de quelqu’un signifie voir les choses du point de vue de cette personne et éprouver des sentiments similaires aux siens. On ne peut jamais ressentir exactement la même chose qu’une autre personne, mais on peut être très proche d’elle grâce à l’empathie. Or, une telle forme d’empathie requiert comme préalable une grande intelligence interpersonnelle, ce dont une IA sera toujours dépourvue à moins qu’elle puisse développer un sentiment d’identité (une singularité dont il est question sur la page https://www.technologyreview.com/s/425733/paul-allen-thesingularity-isnt-near/). Par ailleurs, il faudrait que l’IA soit capable de sentiment, une chose actuellement impossible, et qu’elle soit ouverte au partage des sentiments avec une autre entité (de façon générale, un humain), ce qui est également impossible. En l’état actuel des technologies, une IA ne peut ressentir ni comprendre aucune émotion d’aucune sorte, par conséquent l’empathie est impossible.



Naturellement, la question est de savoir pourquoi l’empathie est si importante. Sans la capacité de ressentir la même chose que quelqu’un d’autre, une IA ne peut pas développer la motivation d’exécuter certaines tâches. Vous pouvez ordonner à l’IA d’exécuter une tâche, mais elle n’aura pas sa propre motivation à le faire. Par conséquent, il est des tâches que l’IA ne pourra jamais effectuer, même si la réalisation de ces tâches est un préalable à l’acquisition des aptitudes et de la connaissance requises pour faire preuve d’une intelligence similaire à l’intelligence humaine.

Développer de vraies relations

À partir des données qu'elle collecte, une IA crée une image de vous. À partir de ces données, elle crée ensuite des schémas, et à l'aide d'algorithmes spécifiques, elle produit un résultat qui donne l'impression qu'elle vous connaît (que vous êtes au moins une connaissance pour elle). Cependant, sachant que l'IA n'a ni sensation ni pensée, elle ne peut pas vous reconnaître en tant que personne. Elle peut vous servir, si vous lui donnez des instructions et en supposant que la tâche fait partie de sa liste de fonctions, mais elle ne peut pas éprouver un sentiment pour vous.

Dans une relation, il faut tenir compte à la fois de l'affinité intellectuelle et des sentiments. L'affinité intellectuelle vient souvent d'un avantage partagé entre les deux entités concernées. Malheureusement, il n'existe aucun avantage partagé entre une IA et un humain (ou une autre entité quelconque). L'IA ne fait que traiter des données en utilisant un algorithme particulier. On ne peut pas dire qu'une chose aime une autre chose dès lors qu'une instruction l'oblige à faire cette proclamation. L'attachement affectif suppose le risque de rejet, ce qui implique la conscience de soi.

Changer de point de vue

Les humains changent parfois d'opinion en fonction d'autre chose que les faits. Même si une certaine ligne de conduite semble prudente, un besoin émotionnel peut en rendre une autre préférable. Cependant, une IA n'a pas de préférences. Elle ne peut donc pas faire un autre choix d'action pour une autre raison qu'un changement dans les probabilités, une *contrainte* (une règle qui l'obligerait à

effectuer ce changement) ou l'exigence de produire un résultat aléatoire.

Faire un acte de foi

La *foi* est la croyance qu'une chose est vraie en l'absence de faits susceptibles de soutenir cette croyance. Dans bien des cas, la foi prend la forme de la *confiance*, qui consiste à croire à la sincérité d'une autre personne en l'absence de preuve que cette personne est fiable. Or, ni la foi ni la confiance n'existent dans une IA, et c'est en partie pour cette raison qu'elle ne peut pas extrapoler des connaissances. L'extrapolation procède souvent de l'intuition ou du sentiment que quelque chose est vrai, malgré l'absence de toute donnée à l'appui de cette hypothèse. Sachant que l'IA n'a pas cette capacité, elle ne peut pas avoir d'idée : or, c'est une condition nécessaire pour présenter des schémas de pensée similaires à ceux des humains.

Les exemples abondent d'inventeurs ayant fait un acte de foi pour créer quelque chose de nouveau. Un des meilleurs exemples est celui d'Edison, qui avait fait au moins mille tentatives pour inventer l'ampoule à incandescence. Une IA aurait renoncé après un certain nombre d'essais, sans doute en raison d'une contrainte. Pour voir une liste de personnes ayant connu de formidables succès grâce à des actes de foi, consultez la page `https://www.uky.edu/~eushe2/Pajares/OnFailingG.html`. Ce sont autant d'exemples de succès hors de portée d'une IA, sachant que l'IA ne peut pas voir plus loin que les données qui lui sont fournies en entrée.

À propos des auteurs

John Mueller est auteur indépendant et rédacteur technique. C'est un « mordu » de l'écriture : à ce jour, il a déjà écrit 108 livres et plus de 600 articles, sur des sujets comme les réseaux, l'intelligence artificielle, la gestion des bases de données et la programmation. Les sujets qu'il traite actuellement sont notamment la science des données, l'apprentissage machine et les algorithmes. Grâce à ses compétences techniques et rédactionnelles, il a aidé plus de 70 auteurs à améliorer le contenu de leurs manuscrits. John Mueller a offert ses services à divers magazines, il a exercé des activités de consultant et il participe à des examens de certification. Ne manquez pas de découvrir son blog : http://blog.johnmuellerbooks.com/.

Luca Massaron est un scientifique spécialisé dans les données, et il est directeur de recherches en marketing. Spécialisé dans l'analyse statistique multivariée, l'apprentissage machine et la connaissance du client, il justifie de plus de dix ans d'expérience dans la résolution de problèmes concrets et la création de valeur par la logique, les statistiques, l'extraction de données et les algorithmes. Passionné par tout ce qui touche à l'analyse de données et par la mise en valeur, auprès des spécialistes comme des profanes, des potentialités de découverte de connaissances à partir des données. Préférant la simplicité à la sophistication non nécessaire, il considère que dans toute discipline, d'importants progrès peuvent être réalisés à partir d'une compréhension en termes simples et d'une mise en pratique de l'essentiel.

Les remerciements de John

Je remercie Rebecca, mon épouse. Même si elle n'est plus là, son esprit est dans chaque livre que j'écris et dans chaque mot qui apparaît sur la page. Elle a cru en moi quand personne d'autre n'y croyait.

Russ Mullen mérite des remerciements pour la révision technique dont il s'est chargé. Il a grandement amélioré la pertinence et la profondeur de ce contenu. Russ me fournit toujours les meilleures adresses de sites Internet pour trouver de nouveaux produits et de nouvelles idées. Il est aussi mon garde-fou dans mon travail.

Matt Wagner, mon agent, mérite ma reconnaissance, tout d'abord pour m'avoir permis d'obtenir le contrat, et pour s'être occupé de tous les détails auxquels les auteurs n'accordent généralement pas leur attention. J'apprécie toujours son aide. Il est agréable de savoir qu'il y a quelqu'un qui veut vous aider.

Un certain nombre de personnes ont lu le livre en totalité ou en partie pour m'aider à affiner l'approche, à tester les exemples de codage, et de façon générale, à mettre dans ce livre tout ce que les lecteurs aimeraient y trouver. Ces bénévoles y ont apporté des contributions trop nombreuses pour qu'il soit possible de les mentionner ici. J'apprécie en particulier les efforts déployés par Eva Beattie et Osvaldo Téllez Almirall, qui ont fourni des commentaires généraux, ont lu le livre en entier et se sont consacrés généreusement à ce projet.

Enfin, je tiens à remercier Katie Mohr, Susan Christophersen et le reste de l'équipe éditoriale et du personnel de la production.

Les remerciements de Luca

Un immense merci, tout d’abord à ma famille, à Yukiko et à Amelia, pour leur soutien, leurs sacrifices et leur aimable patience tout au long de ces jours, de ces nuits, de ces semaines et de ces mois que j’ai passés à travailler à ce livre.

Je tiens à remercier toute l’équipe éditoriale et de production de chez Wiley, en particulier Katie Mohr et Susan Christophersen, pour leur grand professionnalisme et leur soutien à tous les stades de la rédaction de ce livre de la série Pour les Nuls.

Sommaire